Mammalian adaptation of an avian influenza A virus involves stepwise changes in NS1

Running title: Functional evolution of the EIV NS1 gene

Chauché C.[a1], Nogales A.[b1], Zhu H.[a], Goldfarb D.[a], Ahmad Shanizza A.I.[a], Gu Q.[a], Parrish C.R[c]., Martínez-Sobrido L.[b], Marshall J.F.[d], and Murcia P.R.[a]#

[a] MRC-University of Glasgow Centre for Virus Research, Glasgow, United Kingdom.

[b] Department of Microbiology and Immunology, University of Rochester, Rochester, New York, USA.

[c] Baker Institute of Animal Health, Department of Microbiology and Immunology, College of Veterinary Medicine, Cornell University, Ithaca, USA.

[d] Equine Clinical Sciences Division, Weipers Centre Equine Hospital, School of Veterinary Medicine, University of Glasgow, United Kingdom.

[1]C.C and A.N. contributed equally to this article.

To whom correspondence should be addressed:
Pablo Murcia
MRC - University of Glasgow Centre for Virus Research
Institute of Infection, Immunity and Inflammation
College of Medical, Veterinary and Life Sciences
University of Glasgow
Glasgow G61 1QH, UK
Pablo.Murcia@glasgow.ac.uk

ABSTRACT

Influenza A viruses (IAVs) are common pathogens of birds that occasionally establish endemic infections in mammals. The processes and mechanisms that result in IAV mammalian adaptation are poorly understood. The viral non-structural 1 (NS1) protein counteracts the interferon (IFN) response, a central component of the host-species barrier.

We characterised the NS1 proteins of equine influenza virus (EIV), a mammalian IAV lineage of avian origin. We showed that evolutionary distinct NS1s counteract the IFN response using different and mutually exclusive mechanisms: while the NS1s of early EIVs block general gene expression by binding to the cellular polyadenylation specific factor 30 (CPSF30), NS1s from more evolved EIVs specifically block the induction of IFN-stimulated genes by interfering with the JAK/STAT pathway. These contrasting anti-IFN strategies are associated with two mutations that appeared sequentially and became rapidly selected during EIV evolution, highlighting the importance of evolutionary processes on immune evasion mechanisms during IAV adaptation. [150 words]

IMPORTANCE

Influenza A viruses (IAVs) infect certain avian reservoir species, and occasionally transfer to and cause epidemics of infections in some mammalian hosts. However, the processes by which IAVs gain the ability to efficiently infect and transmit in mammals remains unclear. H3N8 equine influenza virus (EIV) is an avian-origin virus that has successfully established a new lineage in horses in the early 1960, and is currently circulating worldwide in the equine population. Here we analysed the molecular evolution of the virulence factor non-structural protein 1 (NS1) and show

that NS1s from different time periods after EIV emergence counteract the host innate immune response using contrasting strategies, which are associated with two mutations that appeared sequentially during EIV evolution. The results shown here indicate that the interplay between virus evolution and immune evasion plays a key role in IAV mammalian adaptation. [139 words]

INTRODUCTION

Influenza A viruses (IAVs) have caused several epizooties in various animal species and four pandemics in humans in the last hundred years. The main natural reservoir of IAVs is wild water birds of various types, but these viruses also circulate in some mammalian host populations – including humans, pigs, horses and dogs (1). However, most infections of mammals by avian IAVs result in either spillover infections or isolated outbreaks (1) (2) and very few have resulted in the establishment of novel endemic lineages. The reasons underlying the establishment of an avian-origin IAV in a mammalian host are still only partially understood.

In most cases it appears that emerging IAVs have to overcome multiple barriers to infect and become established in a new host. For example, the presence of particular forms or linkages of sialic acids within the respiratory tract can facilitate or impair binding by the viral hemagglutinin (HA) or the activity of the viral neuraminidase (NA) (3) (4). Other incompatibilities between IAVs and host proteins can also account for the species barrier as it has been recently shown that ANP32A is a host factor that is required for optimal function of the IAV polymerase complex (5). Molecular incompatibilities can sometimes be overcome by adaptive mutations in the virus, and the high mutation rates exhibited by IAVs can facilitate their appearance. The role of mutations in HA in virus-host interaction are among the best

77 characterised (6), but mutations in other viral segments including those in PB2, NP,
78 NA, M, or NS have also been described (7) (8) (9). However, the roles of the latter in
79 post-transfer adaptation are still incompletely understood.
80 IAV non-structural protein 1 (NS1) is encoded by segment eight and possesses two
81 functionally distinct domains: a N-terminal RNA binding domain (RBD) (amino acids
82 1-73), and a C-terminal effector domain (ED) (85-end) separated by a short and
83 flexible linker region (10) (11). The last 25 residues of NS1 are thought to form a
84 disordered and flexible “tail” (10). Molecular interactions with multiple host proteins
85 allow NS1 to perform a remarkable number of activities (12). One of the best-
86 described functions of NS1 is its ability to antagonise the innate immune response,
87 predominantly the production of type I interferons (IFN). The type I IFN response is
88 an essential component of the host barrier that plays a critical role against emerging
89 viruses, as the recipient hosts usually lack pre-existing immunity. A variety of
90 pathogen recognition receptors (PRRs) recognise pathogen-associated molecular
91 patterns (PAMPs) and activate a cascade of events that lead to the production and
92 secretion of type I IFN (13). Secreted type I IFN binds to its receptor in both infected
93 and neighboring cells and activates the Janus kinase/signal transducers and
94 activators of transcription (JAK/STAT) pathway, which in turn results in the assembly
95 of a protein complex, referred to as interferon-stimulated gene factor 3 (ISGF3),
96 which is composed of phospho-STAT1, phospho-STAT2 and interferon regulatory
97 factor 9 (IRF9). This complex will translocate to the nucleus, bind to DNA regulatory
98 sequences containing IFN-stimulated response elements (ISREs), and stimulate the
99 transcription of hundreds of IFN-stimulated genes (ISGs) including interferon-
100 stimulated gene 15 (ISG15), myxovirus resistance protein 1 (MX1), and 2’-5’-
101 oligoadenylate synthetase 1/2/3 (OAS1/2/3). As a result, IFN-stimulated cells will

102 establish an antiviral state to protect against different viruses, including IAVs (13)
103 (14).
104 H3N8 equine influenza virus (EIV) is an avian-origin virus lineage that has been
105 circulating in horses since at least 1963 (15) (16) and thus provides a model for the
106 long-term mammalian adaptation of avian-derived IAVs. Phylogenetic studies have
107 shown that several amino acid substitutions occurred within each genomic segment
108 during the evolutionary history of EIV, some of which have been associated with host
109 adaptation in other IAVs (17). However, the role of any of those mutations in EIV
110 adaptation to horses is unclear. We hypothesized that the evolution of NS1 would be
111 part of the adaptive process of EIV to horses. To test this hypothesis, we
112 characterized the NS1 genes of a group of evolutionary distinct EIVs using a
113 combination of approaches that included experimental infections, reverse genetics,
114 site-directed mutagenesis, phylogenetics and transcriptomics.
115
116 **RESULTS**
117 **Evolutionary distinct NS1 proteins of the H3N8 EIV lineage exhibit marked**
118 **functional differences.** To experimentally study the functional evolution of the
119 H3N8 EIV NS1 proteins, we selected 13 H3N8 EIVs that included at least one
120 representative virus per decade since its first isolation in 1963 (Table 1) and cloned
121 the coding sequence of their NS1 proteins into a pCAGGS expression plasmid.
122 As the F2/F3 region of CPSF30 involved in interaction with influenza NS1 proteins is
123 conserved between human and equine species (data not shown) and the
124 transfection efficiency of 293T cells is highly superior as the one of E-derm cells
125 (data not shown), we used 293T cells to study the functional evolution of EIV NS1.
126 To compare the ability of each individual NS1 protein to limit the production of IFN-β

we co-transfected 293T cells with NS1 expression plasmids or with an empty vector (pCAGGS), together with a plasmid expressing *Firefly* luciferase under the IFN-β promoter (pIFN-β-FF-Luc), and with a constitutively active *Renilla* luciferase (pREN-Luc) expression plasmid to normalize for the transfection efficacy. At 24 hpt, cells were infected with SeV, a well-described IFN-β inducer. As expected, SeV infection resulted in a strong activation of the IFN-β promoter in cells co-transfected with the empty vector (Figure 1A) (set to 100%) compared to uninfected control cells. However, SeV activation of the IFN-β promoter was blocked in cells transfected with the different EIV NS1 expressing plasmids used in this study (Figure 1A).

We also compared the ability of these NS1 proteins to block the induction of ISGs upon IFN treatment. To this end, we co-transfected 293T cells with the NS1 expression plasmids described above together with a plasmid expressing the *Firefly* luciferase under the control of a promoter containing an IFN-stimulated regulatory element (pISRE-FF-Luc), and again used pREN-Luc as an internal control. At 24 hpt, we treated the cells with uIFN (500 units/well) to stimulate the ISRE-containing promoter and 18 h later we measured FF-Luc and REN-Luc activities. The NS1 proteins of EIVs isolated in 1963 (so called 'Early': U/63 and M/63) displayed a relatively low repression of the ISRE-containing promoter (Figure 1B), and the antagonistic property of NS1 over this promoter increased in EIVs that circulated between 1969 and 1995 (SP/69, F/79, S/89, K/91, LP/95, K/95). Furthermore, the NS1 proteins of EIVs isolated between 1999 and 2013 (so called 'Late': K/99, K/02, O/03, N/03, M/13) showed a variable control of the ISRE-containing promoter depending on the NS1 protein tested (Figure 1B).

To assess the ability of the NS1 proteins to inhibit general gene expression we performed co-transfections of the individual NS1-expression plasmids together with

pREN-Luc and measured luciferase activity 24 h later. The NS1 proteins of EIVs that circulated in the 1960s strongly blocked the otherwise constitutively active *Renilla* luciferase (Figure 1C). In contrast, the NS1 proteins of EIVs isolated after 1979 did not block general gene expression as REN-Luc expression wasn't strongly reduced. In addition, in some cases NS1 expression was associated with an increase in luciferase expression (Figure 1C), as described for other NS1 proteins (18). We also examined the expression levels of NS1 by Western blot in parallel transfected cell lysates. As expected, the NS1 proteins that strongly block general gene expression (i.e. U/63, M/63, SP/69) could not be detected, probably because of the ability of these NS1s to inhibit their own synthesis (19, 20) (Figure 1D). On the other hand, NS1 proteins that did not block host gene expression (F/79 to M/13) were easily detected. Interestingly, the level of expression of NS1 proteins of viruses isolated between 1979 and 1995 (F/79 to LP/95) was lower than that of more recent NS1 proteins (K/95 to M/13).

Taken together, these results suggest that the NS1 proteins of the H3N8 EIV lineage maintained a strong control over IFN-β production throughout evolution, while progressively increasing their control of IFN-stimulated signal transduction. Such control mechanisms are likely independent of the NS1-mediated control of general gene expression, as the NS1 proteins of EIVs that circulated after 1969 lost that ability. These results suggest a direct effect of these NS1 proteins on IFN signaling pathway.

Amino acid 186 and the C-terminal tail affect EIV NS1 protein function. To identify the mutations responsible for the stepwise chronological changes in EIV NS1 function we first examined a multiple sequence alignment of the NS1 proteins used in our studies (Figure 1E). While various amino acid changes were observed, the

most evident was an 11-amino acid truncation in the carboxy-terminus (C-terminus) that appeared in 1999 (21), as well as the appearance of two amino acid substitutions in the 1970s - E186K and A112T. The E186K substitution lies within the putative CPSF30-binding domain and notably, this substitution took place at the same time as the loss of repression of general gene expression by EIV NS1 (Figure 1C).

To determine the impact of the E186K substitution and the C-terminal truncation on NS1 function, we introduced mutations in the NS1 gene of A/equine/Ohio/2003 (O/03), a virus isolated after 40 years of continuous EIV circulation in horses. The NS1 protein of O/03 (O/03) is 219 amino acids long (naturally truncated) and possesses a lysine at position 186 (represented as *Late NS1* in Figure 2A). Three NS1 revertant viruses were tested: a K186E substitution (O/03-K186E), a C-terminal extension of 11 amino acids (O/03-230), and the double change of the extended NS1 and the K186E substitution together (O/03-K186E-230) (Figure 2A). The O/03-230 NS1 revertant would therefore represent the Intermediate (Interm.) NS1 proteins of EIVs that circulated between 1979 and 1995 (F/79, S/89, K/91, LP/95, K/95) (Figure 2A), while the NS1 double revertant represents Early NS1 proteins of EIVs that circulated between 1963 and 1969 (U/63, M/63, SP/69) (Figure 2A). To determine the impact of a glutamic acid residue at position 186, we introduced a K186E mutation in the O/03 NS1 (Artificial). However, it should be noted that such revertant has not been detected in nature.

The presence of a glutamic acid at position 186 in O/03-K186E and O/03-K186E-230 was associated with a decreased ability to control the IFN-β promoter (Figure 2B), a decreased capacity to control the activity of the ISRE-containing promoter (Figure 2C), and a strong repression of the constitutively expressed promoter compared to

O/03 and O/03-230 revertant (Figure 2D). Consistent with this result, NS1 proteins harbouring E186 were not detectable by Western blot (Figure 2E).

Since NS1 binding to CPFS30 is known to block general gene expression (12), we performed co-immunoprecipitation assays using an equine CPSF30 and NS1 O/03 and revertant proteins expressed *in vitro*. We confirmed that the presence of E186 determines CPSF30 binding, regardless of the length of NS1 (Figure 3). It should be noted that binding of NS1 to CPSF30 is consistent with the inhibition of NS1 expression upon transfection (Figure 1D and 2D). This is because transcription from the pCAGGS plasmid is driven by a polymerase-II (pol-II)-dependent promoter through the binding to CPSF30.

Overall, these results suggest that the NS1 proteins of early EIVs were able to control general gene expression through CPSF30 binding, and that K186E substitution, which arose in NS1 after approximately 10 years post viral emergence, influenced the control of IFN and ISG induction. Moreover, this substitution seems to have released the block of general gene expression and improved NS1 ability to control IFN-mediated signal transduction.

Mutations in NS1 acquired throughout evolution affect virus replication and cell-to-cell spread. Three individual O/03 revertant viruses were generated and tested for growth in MDCK cells: one carrying the K186E substitution (O/03-K186E), another carrying the C-terminus extension (O/03-230); and another one carrying both changes (O/03-K186E-230). As indicated in Figure 4A, none of the introduced changes modified the NEP amino acid sequence. While the O/03-K186E and O/03-230 viruses exhibited similar growth kinetics as the O/03 (Figure 4B), the double revertant O/03-K186E-230 displayed a faster replication rate until 12 hpi, but never reached a titre as high as the three other viruses over 72 h (Figure 4B).

In equine cells, the O/03 virus showed a significant advantage over all revertant viruses after 24 hpi. Indeed, O/03 peaked at a titre three to four logs higher than the revertants between 24 and 72 hpi (Figure 4B). Interestingly, during the first 16 hpi, no difference in growth kinetics was observed between O/03 and natural revertant viruses (O/03-230 and O/03-K186E-230). More importantly, O/03-K186E-230 displayed the highest replication rate during the first 12 hpi, reaching a titre two logs higher than any other viruses tested. The O/03-K186E revertant, which expressed an NS1 protein that was never isolated in nature, showed a strong attenuation compared to O/03 between 16 and 72 hpi.

When we tested the ability of O/03 and revertant viruses to spread among neighbouring cells by examining their plaque phenotype in MDCK cells, all revertant viruses displayed significantly smaller plaques than O/03 (Figure 4C), and the O/03-K186E-230 revertant was the most affected, showing pinpoint plaques. This suggested that NS1 evolution resulted in a more efficient cell-to-cell spread.

As NS1 proteins harbouring E186 bind CPSF30 we wanted to check if during viral infection NS1 would be present in the same intracellular compartment as CPSF30 (i.e. the nucleus). All proteins exhibited similar nucleocytoplasmic localisation patterns (Figure 4D), which suggests that the introduced mutations did not affect NS1 sub-cellular localization.

NS1 amino acid 186 and the length of the C-terminal tail are key determinants of EIV control of cellular protein synthesis and the establishment of an antiviral state in equine cells. To compare the viruses' ability to limit the establishment of an antiviral state we infected equine cells (MOI 0.1) with each of the viruses and monitored the expression of two ISGs (ISG15 and MX1) by Western blot at different hpi. As expected, ISG15 and MX1 were detected in lysates of cells

252 infected with both the O/03-230 and O/03-K186E-230 revertant viruses, but not the
253 O/03 (Figure 5A) at 24 hpi. We did not detect those proteins in cells infected with
254 O/03-K186E.
255 To determine the effects of the NS1 revertants on the synthesis of cellular proteins,
256 we performed puromycin assays at various hpi in equine cells (22). Puromycin will
257 be incorporated at the C-terminus of all nascent proteins (23). After cell lysis, an
258 immunoblotting against puromycin is performed to allow the detection of nascent
259 proteins of various sizes (appearing as a black smear) that were being at the
260 moment of cell lysis. The darker and longer the smear is, the more proteins are
261 being produced at the moment of cell lysis. The double revertant virus (O/03-K186E-
262 230) induced a strong protein shutoff (Figure 5B), particularly at early times post
263 infection, and that was associated with high expression levels of NS1 from 8 to 24
264 hpi (Figure 5A). As expected, the presence of E186 did not lead to self-inhibition
265 upon infection (in contrast with the self-inhibition of NS1 observed in transfections,
266 see Figure 3). This is because during viral infections the viral polymerase complex
267 drives transcription of NS1 and this process does not require CPSF30. Differences in
268 molecular mechanisms between transfections and infections have also been
269 previously reported in the literature (18, 24, 25). While this result is consistent with
270 E186 playing an important role in blocking general gene expression via a CPSF30-
271 dependent mechanism, it also suggests that a complementary role of NS1 C-
272 terminal tail might be necessary to induce a strong protein shutoff, as O/03-K186E
273 did not block protein synthesis at any time pi (Figure 5B). However, we cannot rule
274 out the possibility that this lack of protein shutoff was due to a low infection level or
275 low expression of O/03-K186E NS1 protein in equine cells, likely due to a strong
276 attenuation of this virus in equine cells (Figure 6A).

277 Interestingly, the O/03 virus induced a significant increase of protein synthesis at 12
278 hpi, and a low but significant protein shutoff at 24 hpi (Figure 5B).
279 As the double revertant virus (O/03-K186E-230) induced a strong protein shutdown
280 and an early antiviral state, we assumed that the cellular homeostasis would be
281 compromised and the cells would undergo apoptosis prematurely. To test this
282 hypothesis, we compared the expression levels of caspase 3 (cleaved and total) in
283 O/03 and revertant virus-infected E-derm cells (MOI 0.1) at different times pi (Figure
284 5A). As expected, cleavage of caspase 3 was already detectable at 24 hpi in cells
285 infected with the O/03-K186E-230 revertant virus. Interestingly, the O/03-230
286 revertant virus also induced premature apoptosis, suggesting that the tail of NS1
287 plays a role in controlling programmed cell death.
288 **EIV O/03 grows to high levels despite eliciting high amounts of antiviral**
289 **cytokines and can replicate in IFN-primed equine cells.** To assess whether the
290 IFN system was the limiting factor of O/03 revertant virus replication, we measured
291 the antiviral activity of the supernatant of equine cells infected with the viruses (MOI
292 0.1) at different times pi using a rVSV-GFP-based bioassay (26) and compared it
293 with the growth kinetics of each virus (Figure 6). Surprisingly, the highest level of
294 antiviral cytokines was produced upon O/03 infection, peaking at 24 hpi (Figure 6A).
295 This did not seem to affect viral growth since this virus reached rapidly a very high
296 titre (10^7 PFU/ml at 24 hpi), as previously shown (Figure 4B). The pattern of antiviral
297 cytokine production compared to viral growth was similar between O/03-230 and
298 O/03-K186E-230 revertant viruses (Figures 6C and D, respectively), where low
299 levels of antiviral cytokines were associated with limited virus growth. In contrast,
300 cells infected with the O/03-K186E revertant virus exhibited a significantly high peak
301 of antiviral cytokines at 8 hpi that seemed to control virus growth effectively (Figure

6B), since we couldn't detect any viral growth at later times pi. Consistent with this, the level of antiviral cytokines decreased progressively and reached a plateau at 16 hpi (Figure 6B).

As the O/03 virus (but not the NS1 revertants) was able to replicate to high levels in the presence of antiviral cytokines we wanted to determine if it was able to replicate in cells in an antiviral state induced by exogenous IFN. E-derm cells were pretreated with 500U of uIFN for 24 h, and then inoculated with O/03 and NS1 revertant viruses (MOI 0.1). While all the viruses exhibited a reduction in virus growth compared to their observed replication ability in untreated E-derm cells, the O/03 virus reached a titre of 10^5 PFU/ml at 24 hpi, significantly higher than the three other viruses (Figure 7A). These results show that the evolution of NS1 resulted in an increased ability to replicate in cells that have been exposed to uIFN.

To confirm the role of IFN in limiting the replication of the NS1 revertant viruses we treated equine cells 24h prior to infection (MOI 0.1) with Ruxolitinib to block the JAK/STAT pathway and thus impaired the ability of the cells to respond to IFN (27). We compared the growth kinetics of each virus (Figure 7B) in the presence of Ruxolitinib. Although the O/03 virus maintained a significantly higher titre than the three other viruses at 24 and 48 hpi, the revertant viruses reached a higher titre than in normal condition (Figure 4B), confirming the central role of the IFN response in limiting the replication of O/03 NS1 revertant viruses in equine cells.

The length of NS1 and the nature of residue 186 impacts on EIV-mediated control of ISG transcription and general gene expression. To identify further effects of these adaptive mutations on virus-host interactions, we compared the transcriptomes of equine cells infected with the O/03 and revertant viruses. We first determined the total number of differentially expressed genes (DEGs) compared to

mock infected samples (Figure 8) at 8 hpi (end of eclipse phase in E-derm), in order to avoid any bias due to different replication efficiency between viruses. Cells infected with the O/03 virus exhibited the largest number of DEGs to mock-infected cells (n=429), which correlates with an increased production of proteins at 12 hpi (Figure 5B), and likely reflects productive infection. Furthermore, the number of DEGs in revertant virus-infected cells compared to mock-infected samples decreased as follows: 241 for O/03-230-infected samples, 193 for O/03-K186E-infected samples, and 158 for O/03-K186E-230-infected samples (Figure 8A, a detailed list of DEGs including Gene ID, gene name, and $\log_2$ fold change is shown in Table S1). This indicates that these two natural mutations in NS1 have a high impact on the response of equine cells to EIV infection, highlighting their role in virus adaptation.

As the O/03 virus replicated to high titres despite inducing high levels of antiviral cytokines, we compared the number of ISGs and interferon-induced elements that were up-regulated in cells infected with the revertant viruses, but absent in O/03-infected samples (Figure 8B). As expected, a large number of ISGs (i.e. ISG15, MX1, OAS [1/2/3]) that were maintained at a basal level in cells infected with O/03 were up-regulated in cells infected with the revertant viruses (Figure 8B).

We also looked for variations in viral gene expression between O/03 and revertant viruses and did not observe any significant differences (data not shown).

EIV NS1 evolved to block the induction of ISGs at a pre-transcriptional level. Our reporter assays showed that the NS1 proteins that did not bind CPSF30 were able to control ISRE-containing promoters more efficiently (Figure 1 to 3). The transcriptomics data showed that cells infected with O/03 (but not the revertants) exhibited a significant control of ISGs transcripts (Figure 8B) despite inducing high

amounts of antiviral cytokines (Figure 6A). Furthermore, experiments using Ruxolitinib treatment confirmed the central role of the JAK/STAT pathway in limiting the revertant virus growth in equine cells (Figure 7B). Taken together, these results suggested that the NS1 of O/03 virus was blocking the induction of ISGs at a pre-transcriptional level and this involved the JAK/STAT pathway. Since the JAK/STAT signalling pathway, upon IFN stimulation, leads to the transcription of ISGs and this involves the nuclear translocation of the ISGF3 complex. As a marker of ISGF3 nuclear translocation, we compared the amount of nuclear STAT1 at 8 and 24 hpi in equine cells infected with O/03 and revertant viruses. As expected O/03 infection did not result in STAT1 nuclear localization, despite the presence of high levels of antiviral cytokines (i.e. 24 hpi, Figure 6A). In addition, residue 186 played a central role in limiting STAT1 nuclear translocation as O/03-K186E virus infection resulted in abundant nuclear STAT1 (Figure 9A). STAT1 also remained in the cytoplasm of cells infected with O/03-230 or O/03-K186E-230 revertant virus infection, probably due to the low level of antiviral cytokines produced (Figure 6C and D). However, we cannot rule out a possible contribution of NS1 C-terminal tail in the control of STAT1 nuclear translocation. Furthermore, the observed differences were not due to different levels of expression of STAT1, as no significant differences in total STAT1 expression in cell lysates were detected by Western blot (Figure 9B). Finally, the percentage of infected E-derm at 8 hpi was comparable between O/03 and revertant viruses (data not shown), and a higher percentage of cells infected with O/03 was found at 24 hpi, further confirming the significant advantage of O/03 over all revertant viruses after 24 hpi.

DISCUSSION

We investigated the role of NS1 evolution in the post-transfer adaptation of an avian-origin IAV to a mammalian host, focusing on the roles of two mutations that appeared sequentially during EIV evolution. The E186K appeared early after EIV emergence, while the 11-amino acid C-terminal truncation appeared ~20 years later. Our results showed that isogenic viruses carrying the evolved version of NS1 exhibit significant higher fitness in equine cells. The O/03 virus from 2003 was able to replicate to high titres despite high levels of cytokines being produced by infected cells (Figure 6A). This was achieved by blocking the induction of ISGs at a pre-transcriptional level (Figure 8B), which is consistent with the lack of STAT1 nuclear localisation in infected cells (Figure 9A). The ability to replicate in the presence of IFN is a significant fitness trait as it not only renders an important arm of the host antiviral response ineffective, but also acts as a barrier to other respiratory viruses that might compete for the same ecological niche (i.e. the epithelium of the respiratory tract). Co-infection experiments will shed light on this issue.

In contrast, an isogenic virus carrying a revertant version of NS1 (O/03-K186E-230) that mimics the early, more avian-like NS1 proteins of EIV replicated to significantly lower titres over a 72-hour period in equine cells (Figure 4B). This virus induced a cellular protein shutdown (Figure 5B), high levels of apoptosis (Figure 5A), and this was associated with inefficient cell-to-cell spread (Figure 4C). However, O/03-K186E-230 efficiently produced infectious particles at 12 hpi (Figure 4B, 6D). This suggests that the emerging, non-adapted virus was highly susceptible to the host IFN response, and therefore relied on blocking general gene expression non-specifically to transiently control the innate immune response. However, this will likely affect the cellular homeostasis and lead to premature apoptosis, depriving the virus of the cellular resources it needs to replicate. Thus, high replication efficiency

during this early period (Figure 4B) would enhance the chances of onward transmission.

Early EIV NS1 proteins blocked general gene expression by binding to CPSF30, a host protein that plays a central role in pre-mRNA processing (12). A group of residues centred around residue 186 (184 to 188) (25, 28, 29), as well as other additional residues (103, 106, 108, 125) (19, 20, 30) have been previously shown to be involved in this function. Interestingly, for EIV that binding ability was lost with the introduction of K186, a substitution that must have been selectively advantageous for the virus. This is supported by the fact that it became quickly fixed at the virus population scale (Figure 10), and because the E186 in a contemporary EIV NS1 resulted in severe restriction of viral replication (Figure 4B and 6B) and cell-to-cell spread (Figure 4C). That single mutation impaired the ability of EIV to block the induction of the IFN response (Figure 9A), resulting the release of cellular cytokines (Figure 6B), and subsequent up-regulation of ISGs (Figure 8B). This results are consistent with a work we recently published on the evolution of H3N8 canine influenza virus NS1 protein, a virus that originated from the H3N8 EIV lineage in the early 2000s (31).

The second evolutionary change in NS1 that became rapidly fixed at the population level was a C-terminal truncation of the protein that appeared in the mid 1990's (Figure 10). While determining the molecular mechanism underpinning the effect of the C-terminal tail was beyond the scope of this study, our experimental infections suggested that this truncation also increased viral fitness. Indeed, extending the tail of NS1 reduced virus replication (Figure 4B and 6C) and plaque size (Figure 4C), increased viral susceptibility to IFN (Figure 7A and 7B), reduced viral control of ISG transcription (Figure 8B) and expression of ISG15 and MX1 (Figure 5A), and also

resulted in an early induction of apoptosis in infected cells (Figure 5A). The later could be caused by the presence of a putative PDZ-binding domain in the tail of NS1 that has been previously associated with apoptosis (32) (33). Truncation of NS1 would therefore be beneficial for EIV, as it would provide a better control of the IFN response and extend the lifespan of infected cells.

It is important to note that E186 is present in 98.2% of NS1 unique sequences derived from avian, human, swine, canine and equine viruses (n=12775, not shown). However, K186 is more common among equine and canine IAVs. With regard to NS1 length, the majority (63.7 %) of NS1 proteins are longer than 220 amino acids, and such "long" NS1s are highly prevalent among IAVs derived from birds, humans, dogs and horses, but not from pigs as 77% of unique swine NS1 sequences are shorter than 220 residues. Further experimental work is required to unveil the role of residue 186 and the C-terminal length of NS1 in the adaptation of IAVs to other species.

In summary, we have shown that mutations in the NS1 gene that became fixed during the continuous circulation of EIV in horses led to at least two temporally distinct changes in NS1 function that resulted in a better control of the mammalian host innate immune response and likely contributed to the adaptation of an avian-origin virus to its new host, the horse. Comparisons of other mammalian IAV lineages of avian origin will show whether such dynamic strategies are common features of influenza A viruses while adapting to mammals.

MATERIALS AND METHODS

Cells. Madin-Darby Canine Kidney (MDCK: ATCC CCL-34) and Human Embryonic Kidney (293T; ATCC CRL-11268) cells were grown at 37°C, 5% CO_2 in Dulbecco's

452 modified Eagle's medium (DMEM) high glucose, GlutaMax and pyruvate
453 (ThermoFisher Scientific) supplemented with 10% of Fetal Calf Serum (Gibco Life
454 Technologies), and 1% PS (penicillin, 100 units/ml; streptomycin, 100 µg/ml; Gibco
455 Life Technologies). Equine Dermal fibroblasts (E-Derm; ATCC CCL-57) were grown
456 at 37°C, 5% CO_2 in DMEM high glucose, GlutaMax and pyruvate supplemented,
457 15% of Fetal Calf Serum (Gibco Life Technologies), 1% Non-Essential Amino acids
458 (Gibco, Life Technologies), 1% PS (Gibco Life Technologies).
459 **Viruses**. Viral stocks (accession number, abbreviation) of A/equine/Uruguay/1/1963
460 (ACD85423, U/63), A/equine/Miami/1/1963 (ABY81497, M/63),
461 A/equine/SaoPaulo/1/1969 (ACD85390, SP/69), A/equine/Fontainebleau/1/1979
462 (ACD85401, F/79), A/equine/Sussex/1/1989 (ACD97430, S/89),
463 A/equine/Kentucky/1/1991 (ACA24672, K/91), A/equine/LaPlata/1995 (MF182460,
464 LP/95), A/equine/Kentucky/1995 (MF182451, K/95), A/equine/Kentucky/1999
465 (MF182443, K/99), A/equine/Kentucky/5/2002 (ABA42429, K/02),
466 A/equine/Newmarket/5/2003 (ACI48802, N/03), A/equine/Ohio/1/2003 (ABA42431,
467 O/03), A/equine/Mongolia/3/2013 (MF182459, M/13) were grown at passage 2 in
468 MDCK cells (MOI 0.1), then aliquoted and stored at -80^{o}C.
469 Reverse genetic influenza A/equine/Ohio/1/03 (O/03) H3N8 O/03 and NS1 (O/03-
470 K186E, O/03-230 and O/03-K186E-230) revertant viruses were grown in MDCK cells
471 at 37°C, 5% CO_2. For infections, virus stocks were diluted in infection medium
472 (DMEM, 0.3% BSA, 1% PS, and 1 µg/ml of tosylsulfonyl phenylalanyl chloromethyl
473 ketone (TPCK)-treated trypsin (Sigma) (34).
474 Sendai virus (SeV), Cantell strain, was purchased from Charles River Laboratories
475 and stored at -80°C. The rVSV-GFP virus stock used in the bioassay assay was
476 generated by transiently transfecting HEK293T cells in Opti-MEM with an expression

plasmid encoding the VSV glycoprotein of surface (pVSV-G) (26) using TransIT-LT1 transfection reagent (Cambridge Biosciences) at 37°C, 5% CO_2 for 36h (100mm sterile dish format, 5 x 10^6 cells). At the end of the transfection period, the cells were infected with a previous stock of rVSV-GFP virus (2.5 x 10^2 $TCID_{50}$) for 4h at 37°C, 5% CO_2 in growth medium, then washed with PBS and maintained at 37°C, 5% CO_2 in growth medium for a further 16h. The supernatant was then collected and filtered through a 0.45μm filter (Fisher Scientific), before being aliquoted and stored at -80°C.

Cloning of NS1. Mammalian expression constructs for untagged NS1 used in the reporter assay were generated as previously described (18) (19). pcDNA3 plasmids encoding the NS1 were also generated by subcloning from pCAGGS (18). The sequence of each pCAGGS- or pcDNA3-NS1 inserts were confirmed by Sanger sequencing and compared to those available in the NCBI Influenza Database. Viral (v)RNA was extracted from the corresponding viral stocks using the QIAmp Viral RNA kit (Qiagen), and RNA quality was assessed with NanoDrop™ 2000 Spectrophotometer (ThermoFisher Scientific) and stored at -80°C. Reverse transcription (RT) was done using Uni12 primer (35), 500ng of RNA, and SuperScript III Reverse Transcriptase (Invitrogen) following the manufacturer's protocol. Polymerase Chain Reaction (PCR) was done with 100ng of cDNA, PfuUltra II fusion HS DNA polymerase (Agilent) following the manufacturer's protocol, and a specific set of primers for each NS1 was used. Cycling parameters and primers are available upon request.

In order to prevent expression of nuclear export protein (NEP) in the pCAGGS-NS1 plasmids, a silent mutation in the splice acceptor site of NS1 (SAM, splicing acceptor mutation) was introduced for each construct as previously described in (18) (19).

To engineer the three NS1 revertants, Lysine (K)-to-Glutamic acid (E) (codon 186 'AAA' changed into 'GAA') and Stop Codon-to-Arginine (R) (codon 220 'TGA' replaced by 'CGA'), mutations were introduced into the pCAGGS-splice acceptor revertant O/03 NS1 constructs and the ambisense pDP2002 plasmid encoding O/03 O/03 NS gene using the PFUTurbo DNA polymerase (Agilent) and specific sets of primers. The presence of introduced mutations was confirmed by Sanger sequencing. Cycling parameters and primers are available upon request.

Reporter assay. For analysis of IFN-β and ISRE promoter activation and general gene expression, 293T cells (12-well plate format, $2.5x10^5$ cells/well) were transiently co-transfected using TransIT-LT1 transfection reagent (Cambridge Biosciences), with either 50ng of pIFN-β-FF-Luc or 50ng of pISRE-FF-Luc (reporter plasmids encoding Firefly luciferase [FF-Luc] under the control of the IFN-β promoter or the ISRE promoter, respectively), 50ng of a plasmid constitutively expressing Renilla luciferase (pREN-Luc) under the SV40 promoter (26) (kindly provided by Benjamin G. Hale), as well as 1000ng of the indicated pCAGGS-SAM NS1 expressing plasmids (or empty pCAGGS,1000ng). After 24 h of transfection, cells were infected with 50 hemagglutinating units (HAU) of SeV for 18 h, then lysed with 250µl of passive lysis buffer (Promega). IFN- β-FF-Luc, ISRE-FF-Luc and REN-Luc activities were measured using the Dual-luciferase reporter assay system (Promega), as directed by the manufacturer's protocol. All transfections were carried out in triplicates, and experiments were repeated independently three times.

Co-immunoprecipitation of NS1 with CPSF30. The equine CPSF30 was cloned as previously described for the human CPSF30 (18). Briefly, the equine CPSF30 gene was amplified by RT-PCR from equine cells using oligo d(T) and PCR with specific primers. The RT-PCR product was then cloned into the pCAGGS HA-COOH plasmid

as described previously (36). NS1 proteins were synthesized *in vitro* using pcDNA3 plasmids and the TNT7 transcription/translation kit (Promega) following the manufacturer's recommendations. Human 293T cells (1.5×10^6 cells/well, 6-well format, triplicates) were transiently transfected with 2 µg/well of an HA-equine CPSF30-expressing pCAGGS plasmid. Fourty-eight hours post-transfection (hpt), cells were lysed in 20 mM Tris-HCl (pH 7.5), 100 mM NaCl, 0.5 mM EDTA, 5% glycerol, 1% Triton X-100 supplemented with a complete mini protease inhibitor cocktail (Pierce). Cleared cell lysates were incubated overnight at 4°C with the *in vitro*-synthesized O/03 and revertant NS1 proteins and 20 µl of an anti-HA affinity resin (Sigma). After extensive washing, precipitated proteins were dissociated from the resin using Laemmli buffer + β-mercaptoethanol. Total proteins from cells lysates were then analyzed by SDS-PAGE and Western blotting as described below, using a primary specific rabbit polyclonal antibody against NS1 [Genscript] or HA-tag for CPSF30 (Sigma).

Rescue of recombinant H3N8 EIVs. Viruses were rescued as previously described (37). Briefly, co-cultures (2:3) of 293T/MDCK cells (6-well plate format, 2×10^6 cells/well) were seeded 24h prior to viral rescue. Cells were transiently co-transfected using TransIT-LT1 (Cambridge Bioscience), with 2.5µg of seven-ambisense O/03 plasmids (pDP2002-PB2, -PB1, -PA, -HA, -NP, -NA, -M) plus ambisense O/03 NS plasmid (O/03) or the NS revertant constructs (O/03-K186E, O/03-230 and O/03-K186E-230). At 24 hpt, the medium was replaced by infection medium. Virus-containing tissue culture supernatants were collected 2 to 3 days pt, clarified, and used to infect fresh MDCK cells (P1 stock) for 2 to 3 days post-infection. Viral titres of P1 stocks were determined by immunofocus assay (focus forming units, FFU/ml) in MDCK cells, using the mouse monoclonal anti-NP antibody

(clone HB-65, European Veterinary Laboratory), the horseradish peroxidase-conjugated rabbit anti-mouse IgG antibody (AbD Serotec) and TrueBlue peroxidase substrate (Insight Biotechnology), as previously described (37). These P1 stocks were then used to grow P2 viral stocks at MOI 0.01 in MDCK for further use in experiments. For experimental infections, a minimum of two viral stocks for each virus were rescued and grown independently. The NS segment of each virus was sequenced by the Sanger method to confirm the sequence of the O/03 or revertant viruses.

Viral infection. Confluent monolayers of MDCK cells (12-well plate format, triplicates, 5×10^5 cells/well) or E-derm cells (12-well plate format, triplicates, 2.5×10^5 cells/wells) were infected (MOI 0.01 and 0.1, respectively) with the indicated viruses and placed at 37°C, 5% CO_2. Cells were grown on coverslips for confocal microscopy. After 1h incubation, cells were washed with PBS and infection medium was replaced with 500 µl of fresh growth medium. Tissue culture supernatants were collected at various times pi and stored at -80°C, and cells were fixed in 0.1% buffered formalin at 4°C for 16h and kept for confocal microscopy or fluorescence-activated cell sorting (FACS) analysis. Each experiment was repeated three times independently. Viral titres were determined by immunofocus assay in MDCK cells. Titrations were repeated three times independently, and the mean value and standard error mean were calculated using GraphPad Prism7.

To measure virus growth kinetics in the presence of an inhibitor of the IFN response, E-derm cells were treated with Ruxolitinib (Selleck Chemicals). Ruxolitinib was prepared as 10mM stocks in dimethyl sulfoxide (DMSO), and used at a concentration of 4µM (27). Treatment was started 24h prior to infection and maintained at the same concentration for the whole experiment.

For plaque phenotype, a confluent monolayer of MDCK cells (6-well plate format, triplicates, 6.4 × 10^4 cells/well) were infected with serial dilution (1:2) of the viral stock of interest and placed at incubation at 37°C, 5% CO2, humidified atmosphere. Plates were gently rocked every 10 minutes. After 1h incubation, cells were gently washed with PBS and infection medium was replaced with a 50:50 2.4% Avicell:2X MEM overlay for a 48h strictly. At 48 hpi, the overlay was discarded, cells were gently washed 3 times with PBS and fixed with 80% ice-cold acetone solution (Sigma-Aldrich) for 10 minutes at room temperature. The plates were then let to dry overnight at room temperature before being treated with 1% Triton (Triton™ X100, Sigma-Aldrich-Aldrich) X100 PBS solution for 10 minutes at room temperature, followed by 1 hour of incubation with 10% NGS PBS solution at room temperature. This was followed by an overnight immunoblotting at 4°C in 10% Normal Goat Serum plus PBS with a monoclonal anti-influenza A virus nucleoprotein (NP) antibody (clone HB65, European Veterinary Laboratory, 1/500 dilution) under gentle agitation. After a 3-step washing with PBS, a horseradish peroxidase-conjugated rabbit anti-mouse IgG antibody (AbD Serotec, UK, 1/1000 dilution) was used in PBS solution for a further 4h at room temperature under gentle agitation. A color development method was used to reveal the immunofocus using the TrueBlue peroxidase substrate (Insight Biotechnology), where 2ml of substrate was used per well for exactly 10 minutes, before being stop with tap water.

Viral and cellular protein staining for FACS and confocal microscopy. Cells were permeabilized with 1% Triton X100 for 10min, and blocked in PBS 10% Normal Goat Serum (Gibco, Life Technologies) for 1h. Cells were then incubated with anti-influenza A NP protein antibody, rabbit polyclonal anti-NS1 protein antibody (Genscript) or rabbit STAT1 polyclonal antibody (Santa Cruz) overnight at 4°C. Cells

were then washed twice with PBS and incubated for 4h with a rabbit anti-mouse IgG Alexa fluor 488 (Cell signalling) or donkey anti-rabbit IgG Alexa fluor 555 (Cell signalling), before FACS analysis (Guava Flow Cytometer, Merck) or fixed in VECTASHIELD Antifade Mounting Medium with DAPI (Vector laboratories) and analysed by confocal microscopy.

For confocal microscopy, images were taken with a 63× oil objective on a Zeiss LSM 880 confocal microscope with GaAsP detector. All images were taken with cross-talk minimized using best signal calculations in Zen software. 3 × 3 tilescans were collected in three different positions for each sample. The experiment was repeated three times independently. Images were imported into ImageJ and STAT1 nuclear localisation was analysed in NP positive cells, with a minimum of 1000 NP-expressing cells were analysed for each sample. The nuclear localisation of STAT1 was analysed using the 'AND' function in the image calculator, and was accomplished by detecting co-localisation signals between STAT1 and DAPI, using the co-localization threshold function in ImageJ. The quantity of nuclear STAT1 (nSTAT1) was expressed as a percentage of total STAT1 (tSTAT1) in infected cells. For the representative pictures of nSTAT1, the signal intensity of nSTAT1 was normalized to O/03-K186E at 8hpi.

SDS-PAGE and Western blotting. Cells were lysed in Laemlli buffer + β-mercaptoethanol and stored immediately at -80°C. Samples were boiled for 15 mins at 95°C prior to polypeptide separation by SDS-PAGE on NuPAGE Novex 4-12% Bis-Tris protein gels (ThermoFisher Scientific). Proteins were detected by Western blotting following to transfer to nitrocellulose membranes. The membranes were blocked for 1h at room temperature in 5% milk TBST 0.1% Tween-20 and immunoblotted overnight at 4°C in 5% milk TBST-0.1% Tween-20 with the following

627 antibodies: mouse monoclonal anti-Mx1 (clone M143, provided by Dr. Georg Kochs,
628 University of Freiburg), rabbit polyclonal anti-ISG15 (Proteintech), rabbit polyclonal
629 anti-NS1 (GenScript), rabbit monoclonal anti-cleaved caspase-3 and rabbit
630 monoclonal anti-caspase-3 (Cell Signalling), rabbit monoclonal anti-y-Tubulin
631 (Sigma), mouse monoclonal anti-puromycin (clone 12D10, Millipore), anti-mouse IgG
632 HRP conjugated (Dako) and anti-rabbit IgG HRP conjugated (Dako). The
633 chemiluminescent signal was detected using Amersham ECL Prime Western Blotting
634 Detection Reagent (GE-Healthcare), and captured with ChemiDoc XRS+ System
635 (Biorad).

636 **Antiviral cytokine production and general protein shutdown**. E-derm cells (12-
637 well plate format, triplicates, 2.5 x 10^5 cells/well) were infected (MOI 0.1) with the
638 indicated viruses for a total of 72 h. At the indicated times post-infection,
639 supernatants were collected and stored at -80°C for further analysis by bioassay,
640 while cells were treated for 1h at 37°C, 5% CO_2 with Puromycin (20µg/ml in DMEM,
641 15% FBS) prior to lysis in Laemlli buffer + β-mercaptoethanol and stored at -80°C for
642 further analysis by western blot. Puromycin is a well-known antibiotic that competes
643 against aminoacyl tRNA on the ribosome A site (38). As such, puromycin enables
644 examination of total protein production without requiring transfection, radio-labeling,
645 or the prior choice of a candidate gene **(39).** If cells are incubated with puromycin,
646 lysed and immunoblotted using an anti-puromycin antibody, all the proteins being
647 produced will be immunostained as puromycin will be incorporated at the C-terminus
648 of all nascent proteins. For the bioassay, supernatants were UV-inactivated for 5
649 minutes at room temperature and used to treat fresh E-derm cells (48-well plate
650 format, 6x10^4 cells/well, triplicates) for 24h. The cells were then infected with rVSV-
651 GFP virus (2.5 x 10^2 $TCID_{50}$) for 8h, then trypsinized and fixed in 0.1% buffered

formalin for 16h at 4°C. The percentage of GFP-expressing cells were then analysed by FACS. For controls, E-derm cells were mock treated or treated with 500 units of universal type I IFN (uIFN; PBL Assay Science). GFP expression of mock-treated cells infected with rVSV-GFP was considered as 100% and GFP expression of uIFN-treated cells infected with rVSV-GFP was considered as 0%. Mean values and standard error means were calculated with GraphPad Prism7.

RNAseq. Confluent monolayers of E-derm cells (12-well plates, triplicates, 2.5×10^5 cells/wells) were infected with O/03 and revertant viruses (MOI 1) or mock infected at least three times independently. At 8 hpi cells were washed with PBS, immunostained with the anti-NP antibody and the proportion of infected cells was determined by FACS. The samples containing a similar proportion of infected cells were selected for transcriptomic analysis, and cells were lysed with 500µl of TRIzol (ThermoFisher Scientific) for further RNA extraction.

Total RNA was extracted using the TRIzol method and further purified using the RNeasy mini spin columns (Qiagen), including an on-column DNase I digestion step (Qiagen) according to the manufacturer's protocol. RNA concentration was measure with Qubit and the Qubit RNA HS Assay Kit (ThermoFisher Scientific) following the manufacturer's protocol. The ribosomal (r)RNA integrity number was measured using an Agilent 2100 BioAnalyzer (Agilent Technologies).

4.5ug of total RNA was enriched by selectively depleting rRNA using the RiboMinus™ Eukaryote Kit v2 (Ambion, Life Technologies). The sequence reads (GenBank accession number: PRJEB21264) were processed according to the Tuxedo pipeline (40). Read quality was assessed using FastQC, and TopHat2 and Bowtie2 were used to map short reads against the Equus caballus 2 genome (GCA_000002305.1). A list of differentially expressed genes to mock-infected

samples was generated using CuffDiff2 (genes with Benjamini Hochberg-p value <0.05 were considered significant) (41).

Phylogenetic analysis of EIV NS1 sequences. We collected 170 EIV H3N8 NS sequences from the NCBI Influenza Virus Resource database (sequences available upon request). We also included four EIV H3N8 NS sequences that we sequenced for this study (Accession numbers: MF182460, MF182451, MF182443, MF182459) (42). SeaView Version 4.6.1 (43) was used to align the NS1 coding regions and the final alignment was edited manually. NEP coding regions were removed from the alignment. We used BEAST Version 1.8.4 (44) to infer maximum clade credibility (MCC) trees. We used a strict molecular clock and HKY85+G model of nucleotide substitution. Each codon position was estimated with separate substitution rates and nucleotide frequencies. Two individual chains were run until convergence was achieved.

Statistical analysis. Unless stated otherwise, significance was calculated by Two-way ANOVA followed by Bonferroni's multiple comparisons test.

[Text word count: 7,301]

ACKNOWLEDGMENTS

We would like to thank Benjamin G. Hale for his input on this project and for providing reagents to carry out this work. The authors declare no competing interests.

This research was partially funded by the Horserace Betting Levy Board (Grant 779) and Veterinary Fund Small Grant Scheme of the University of Glasgow to CC. CC was supported by a Veterinary Research Training Scholarship funded by the

Horserace Betting Levy Board (VET/RS/2512). PRM was supported by the Medical Research Council of the United Kingdom (Grant MC_UU_120/14/9). LM-S and CP were partially supported by the University of Rochester Technology Development Fund and by the New York Influenza Center of Excellence (NYICE, HHSN272201400005C), a member of the NIAID Centers of Excellence for Influenza Research and Surveillance, CEIRS (LM-S).

REFERENCES

1. **Parrish CR, Murcia PR, Holmes EC.** 2015. Influenza virus reservoirs and intermediate hosts: dogs, horses, and new possibilities for influenza virus exposure of humans. Journal of virology **89:**2990-2994.
2. **Webby RJ, Webster RG.** 2001. Emergence of influenza A viruses. Philosophical transactions of the Royal Society of London Series B, Biological sciences **356:**1817-1828.
3. **Gamblin SJ, Skehel JJ.** 2010. Influenza hemagglutinin and neuraminidase membrane glycoproteins. The Journal of biological chemistry **285:**28403-28409.
4. **Wasik BR, Barnard KN, Parrish CR.** 2016. Effects of Sialic Acid Modifications on Virus Binding and Infection. Trends in microbiology **24:**991-1001.
5. **Long JS, Giotis ES, Moncorge O, Frise R, Mistry B, James J, Morisson M, Iqbal M, Vignal A, Skinner MA, Barclay WS.** 2016. Species difference in ANP32A underlies influenza A virus polymerase host restriction. Nature **529:**101-104.
6. **Lipsitch M, Barclay W, Raman R, Russell CJ, Belser JA, Cobey S, Kasson PM, Lloyd-Smith JO, Maurer-Stroh S, Riley S, Beauchemin CA, Bedford T, Friedrich TC, Handel A, Herfst S, Murcia PR, Roche B, Wilke CO, Russell CA.** 2016. Viral factors in influenza pandemic risk assessment. eLife **5**.
7. **Le QM, Sakai-Tagawa Y, Ozawa M, Ito M, Kawaoka Y.** 2009. Selection of H5N1 influenza virus PB2 during replication in humans. Journal of virology **83:**5278-5281.
8. **Sakabe S, Ozawa M, Takano R, Iwastuki-Horimoto K, Kawaoka Y.** 2011. Mutations in PA, NP, and HA of a pandemic (H1N1) 2009 influenza virus contribute to its adaptation to mice. Virus research **158:**124-129.
9. **Dankar SK, Miranda E, Forbes NE, Pelchat M, Tavassoli A, Selman M, Ping J, Jia J, Brown EG.** 2013. Influenza A/Hong Kong/156/1997(H5N1) virus NS1 gene mutations F103L and M106I both increase IFN antagonism, virulence and cytoplasmic localization but differ in binding to RIG-I and CPSF30. Virology journal **10:**243.

10. **Hale BG, Barclay WS, Randall RE, Russell RJ.** 2008. Structure of an avian influenza A virus NS1 protein effector domain. Virology **378:**1-5.
11. **Wang X, Basler CF, Williams BR, Silverman RH, Palese P, Garcia-Sastre A.** 2002. Functional replacement of the carboxy-terminal two-thirds of the influenza A virus NS1 protein with short heterologous dimerization domains. Journal of virology **76:**12951-12962.
12. **Ayllon J, Garcia-Sastre A.** 2015. The NS1 protein: a multitasking virulence factor. Current topics in microbiology and immunology **386:**73-107.
13. **Garcia-Sastre A, Biron CA.** 2006. Type 1 interferons and the virus-host relationship: a lesson in detente. Science **312:**879-882.
14. **Randall RE, Goodbourn S.** 2008. Interferons and viruses: an interplay between induction, signalling, antiviral responses and virus countermeasures. The Journal of general virology **89:**1-47.
15. **Marois P, Pavilanis V, Boudreault A, Di Franco E.** 1963. An Outbreak of Type A(2) Influenza Among Horses. Canadian journal of comparative medicine and veterinary science **27:**257-260.
16. **Scholtens RG, Steele JH, Dowdle WR, Yarbrough WB, Robinson RQ.** 1964. U.S. Epizootic of Equine Influenza, 1963. Public health reports **79:**393-402.
17. **Murcia PR, Wood JL, Holmes EC.** 2011. Genome-scale evolution and phylodynamics of equine H3N8 influenza A virus. Journal of virology **85:**5312-5322.
18. **Kochs G, Garcia-Sastre A, Martinez-Sobrido L.** 2007. Multiple anti-interferon actions of the influenza A virus NS1 protein. Journal of virology **81:**7011-7021.
19. **Hale BG, Steel J, Medina RA, Manicassamy B, Ye J, Hickman D, Hai R, Schmolke M, Lowen AC, Perez DR, Garcia-Sastre A.** 2010. Inefficient control of host gene expression by the 2009 pandemic H1N1 influenza A virus NS1 protein. Journal of virology **84:**6909-6922.
20. **Kochs G, Garcia-Sastre A, Martinez-Sobrido L.** 2007. Multiple anti-interferon actions of the influenza A virus NS1 protein. J Virol **81:**7011-7021.
21. **Quinlivan.** 2005. Attenuation of Equine Influenza Viruses through Truncations of the NS1 Protein. Journal of Virology.
22. **Varela M, Pinto RM, Caporale M, Piras IM, Taggart A, Seehusen F, Hahn K, Janowicz A, de Souza WM, Baumgartner W, Shi X, Palmarini M.** 2016. Mutations in the Schmallenberg Virus Gc Glycoprotein Facilitate Cellular Protein Synthesis Shutoff and Restore Pathogenicity of NSs Deletion Mutants in Mice. Journal of virology **90:**5440-5450.
23. **Tabuchi I.** 2003. Next-generation protein-handling method: puromycin analogue technology. Biochem Biophys Res Commun **305:**1-5.
24. **Hale BG, Steel J, Medina RA, Manicassamy B, Ye J, Hickman D, Hai R, Schmolke M, Lowen AC, Perez DR, Garcia-Sastre A.** 2010. Inefficient control of host gene expression by the 2009 pandemic H1N1 influenza A virus NS1 protein. J Virol **84:**6909-6922.
25. **Nemeroff ME, Barabino SM, Li Y, Keller W, Krug RM.** 1998. Influenza virus NS1 protein interacts with the cellular 30 kDa subunit of CPSF and inhibits 3'end formation of cellular pre-mRNAs. Mol Cell **1:**991-1000.
26. **Martinez-Sobrido L, Zuniga EI, Rosario D, Garcia-Sastre A, de la Torre JC.** 2006. Inhibition of the type I interferon response by the nucleoprotein of

the prototypic arenavirus lymphocytic choriomeningitis virus. Journal of virology **80:**9192-9199.

27. **Stewart CE, Randall RE, Adamson CS.** 2014. Inhibitors of the interferon response enhance virus replication in vitro. PloS one **9:**e112014.
28. **Noah DL, Twu KY, Krug RM.** 2003. Cellular antiviral responses against influenza A virus are countered at the posttranscriptional level by the viral NS1A protein via its binding to a cellular protein required for the 3' end processing of cellular pre-mRNAS. Virology **307:**386-395.
29. **Twu KY, Noah DL, Rao P, Kuo RL, Krug RM.** 2006. The CPSF30 binding site on the NS1A protein of influenza A virus is a potential antiviral target. J Virol **80:**3957-3965.
30. **Das K, Ma LC, Xiao R, Radvansky B, Aramini J, Zhao L, Marklund J, Kuo RL, Twu KY, Arnold E, Krug RM, Montelione GT.** 2008. Structural basis for suppression of a host antiviral response by influenza A virus. Proc Natl Acad Sci U S A **105:**13093-13098.
31. **Nogales A, Chauche C, DeDiego ML, Topham DJ, Parrish CR, Murcia PR, Martinez-Sobrido L.** 2017. Amino acid substitution K186E in the canine influenza virus H3N8 NS1 protein restores its ability to inhibit host gene expression. J Virol doi:10.1128/JVI.00877-17.
32. **Liu H, Golebiewski L, Dow EC, Krug RM, Javier RT, Rice AP.** 2010. The ESEV PDZ-binding motif of the avian influenza A virus NS1 protein protects infected cells from apoptosis by directly targeting Scribble. Journal of virology **84:**11164-11174.
33. **Golebiewski L, Liu H, Javier RT, Rice AP.** 2011. The avian influenza virus NS1 ESEV PDZ binding motif associates with Dlg1 and Scribble to disrupt cellular tight junctions. Journal of virology **85:**10639-10648.
34. **Martinez-Sobrido L, Garcia-Sastre A.** 2010. Generation of recombinant influenza virus from plasmid DNA. Journal of visualized experiments : JoVE doi:10.3791/2057.
35. **Hoffmann E, Stech J, Guan Y, Webster RG, Perez DR.** 2001. Universal primer set for the full-length amplification of all influenza A viruses. Archives of virology **146:**2275-2289.
36. **Martinez-Sobrido L, Giannakas P, Cubitt B, Garcia-Sastre A, de la Torre JC.** 2007. Differential inhibition of type I interferon induction by arenavirus nucleoproteins. Journal of virology **81:**12696-12703.
37. **Gonzalez G, Marshall JF, Morrell J, Robb D, McCauley JW, Perez DR, Parrish CR, Murcia PR.** 2014. Infection and pathogenesis of canine, equine, and human influenza viruses in canine tracheas. Journal of virology **88:**9208-9219.
38. **Monro RE, Cerna J, Marcker KA.** 1968. Ribosome-catalyzed peptidyl transfer: substrate specificity at the P-site. Proc Natl Acad Sci U S A **61:**1042-1049.
39. **Starck SR, Green HM, Alberola-Ila J, Roberts RW.** 2004. A general approach to detect protein expression in vivo using fluorescent puromycin conjugates. Chem Biol **11:**999-1008.
40. **Trapnell C, Hendrickson DG, Sauvageau M, Goff L, Rinn JL, Pachter L.** 2013. Differential analysis of gene regulation at transcript resolution with RNA-seq. Nature biotechnology **31:**46-53.
41. **Ratinier M, Shaw AE, Barry G, Gu Q, Di Gialleonardo L, Janowicz A, Varela M, Randall RE, Caporale M, Palmarini M.** 2016. Bluetongue Virus

NS4 Protein Is an Interferon Antagonist and a Determinant of Virus Virulence. Journal of virology **90:**5427-5439.

42. **Bao Y, Bolotov P, Dernovoy D, Kiryutin B, Zaslavsky L, Tatusova T, Ostell J, Lipman D.** 2008. The influenza virus resource at the National Center for Biotechnology Information. J Virol **82:**596-601.

43. **Gouy M, Guindon S, Gascuel O.** 2010. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol Biol Evol **27:**221-224.

44. **Drummond AJ, Suchard MA, Xie D, Rambaut A.** 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol Biol Evol **29:**1969-1973.

FIGURE LEGENDS

Figure 1. Functional characterization of evolutionary distinct EIV NS1 proteins. A schematic representation of three versions of the NS1 protein (early, intermediate and late) are found throughout EIV evolution. Early NS1 proteins, full length (230 amino acids) harboring E186, were found in EIVs circulating between the 1960s and 1970s; intermediate NS1s, full length with K186, were found in EIVs circulating between the late 1970s and late 1990s; and late NS1s, truncated at their C-terminus (219 amino acids) with K186, are found in EIVs isolated since the late 1990s. **(A-D)** Human 293T cells were transiently co-transfected with a pCAGGS expression plasmid encoding the indicated NS1 proteins (or empty plasmid), together with either **(A)** a *Firefly* luciferase IFN-β promoter reporter construct (pIFN-β-FF-Luc), or **(B)** a plasmid expressing the *Firefly* luciferase under the control of a promoter containing an IFN-stimulated regulatory element (pISRE-FF-Luc), as well as with a constitutively active *Renilla* luciferase (pREN-Luc) expression plasmid. At 24 hpt, cells were either **(A)** infected with SeV (+) or mock infected (-), or **(B)** treated with universal interferon (uIFN) (+) or mock treated (-) for 18h. **(A-B)** The relative activity of the respective FF-Luc activity in cells transfected with (**A**) the pIFN-β-FF-Luc or **(B)** pISRE-FF-Luc plasmids were determined as the ratio between FF-Luc and REN-Luc in each corresponding sample. Values were normalized to **(A)** empty pCAGGS

plasmid (+) SeV or **(B)** empty pCAGGS plasmid (+) uIFN (set to 100%). **(C)** Human 293T cells were transiently co-transfected individually with the indicated NS1-expression plasmids together with pREN-Luc. Total REN-Luc levels were measured 24 h later, and values were normalized to empty pCAGGS transfected cells. **(A-C)** Bars represent means and error bars SEM, from three independent experiments. Significance was calculated as indicated in material and method. ****, p<0.0001 for pCAGGS+SeV or pCAGGS+uIFN versus all other conditions. £ symbols indicate significance between indicated NS1 against all other NS1 proteins, with ££, p<0.01; ££, p<0.001; £££, p<0.0001. **(D)** The level of expression of the indicated NS1 was analyzed by Western blotting from parallel transfected lysates using a rabbit polyclonal anti-EIV NS1 antibody. **(E)** Alignment of NS1 amino acid sequences of each of the 13 EIVs used in this study. Polymorphism between the 13 NS1 proteins are indicated in yellow. Red shadings represent NS1 amino acid changes investigated in the study. Grey shadings show other polymorphisms present at high frequency in the EIV population.

Figure 2. NS1 E186 controls general gene expression.

(A) Protein alignment showing the sequence of interest and the introduced mutations in EIV NS1. Below the alignment there is a diagram of the three different NS1 proteins that represent evolutionary distinct NS1 proteins. An early NS1 (shown in red) is O/03-230 amino acids long and possesses E186; an intermediate NS1 (shown in light brown) is O/03-230 amino acids long and possesses K186; and a Late NS1 (shown in blue) is 219 amino acids long and possesses K186. **(B)** IFN-β, **(C)** ISRE, and **(D)** general gene expression reporter assays were done using the indicated NS1 pCAGGS expression constructs as described in Figure 1.

896 Experiments were performed three times independently and significance was
897 calculated as described in material and method. ****, p<0.0001 for pCAGGS+SeV or
898 pCAGGS+uIFN versus all other conditions. ££££ indicates a significant difference
899 (p<0.0001) between K186E-containing revertant(s) and O/03 and O/03-230
900 revertant. €€€€ indicates significance (p<0.0001) between O/03-K186E and all other
901 conditions, and $$$$ indicates a significant difference (p<0.0001) between O/03 and
902 O/03-230 revertant. Error bars represent SEM. **(E)** Western blotting from lysates
903 from parallel samples was performed as described in Figure 1.
904
905 **Figure 3. Amino acid 186 is a determinant of NS1 interaction with CPSF30.**
906 NS1 variants were synthesized *in vitro*. Human 293T cells were transiently
907 transfected with 2000ng of a pCAGGS plasmid expressing a HA-tagged version of
908 the equine CPSF30. At 30 hpt, cells were lysed and cleared cell lysates expressing
909 HA-CPSF30 were incubated overnight at 4°C with the *in vitro*-synthesized NS1
910 proteins and 20 µl of an anti-HA affinity resin. Precipitated proteins were dissociated
911 from the resin and analyzed by Western blotting. A representative picture of protein
912 expression levels ('input', left panel) and co-immunoprecipitation results (IP anti-HA,
913 right panel) are shown.
914
915 **Figure 4. Characterization of EIVs carrying mutations in NS1.**
916 **(A)** Schematic representation of the mutations introduced in O/03 NS segment at
917 NS1 codon 186 and codon 220 (nucleotides 582 to 584 and 684 to 686 in O/03 NS
918 segment, respectively), and the resulting amino acid changes in the different NS1
919 proteins (late, artificial, intermediate, early) are indicated. A(n) indicates a poly-A tail,
920 and '*' indicates a stop codon (TGA). The introduced changes did not change the

NEP amino acids at positions 28 (TTA/G coding for leucine) and 68 (AAT/C coding for asparagine). **(B)** Growth kinetics of O/03, O/03-K186E, O/03-230, and O/03-K186E-230 viruses in MDCK cells (MOI 0.01) and E-derm cells (MOI 0.1). Data points represent virus titers in supernatants at different times post-infection and error bars represent SEM. **(C)** Plaque phenotype of O/03, O/03-K186E, O/03-230, and O/03-K186E-230 in MDCK cells at 48 hpi. The average plaque size for each virus was determined by counting over 100 plaques per condition. Significance was calculated as described in Methods. +++, $p<0.001$ for O/03-K186E-230 versus O/03. ****, $p<0.0001$; ***, $p<0.001$; **, $p<0.01$ for O/03 versus three other viruses. £, $p<0.05$ for O/03-K186E-230 versus O/03-K186E. **(D)** Subcellular localization of NS1 in infected equine cells. E-derms were infected with O/03, O/03-K186E, O/03-230, and O/03-K186E-230 (MOI 0.1) for 24h. NS1 was detected by immunofluorescence as described in Methods. The experiment was performed three times independently and representative pictures of NS1 are shown. Scale represents 20μm.

Figure 5. NS1 amino acid 186 and C-terminal tail affect EIV control of general protein production, apoptosis and response to IFN in equine cells.

E-derm cells were infected (MOI 0.1) with O/03, O/03-K186E, O/03-230, and O/03-K186E-230 viruses. **(A)** Cells were lysed and immunoblotted for the indicated proteins at indicated hpi. **(B)** To measure protein shutoff upon infection, culture media were replaced with a puromycin-containing suspension and incubated for 1h. The cells were then lysed and immunoblotted for puromycin at indicated hpi, as described in Methods. Western blot quantification for puromycin and γ-tubulin signal were determined for three independent experiments and expressed relatively to mock infected cells. Significance was calculated as indicated in material and method.

946 ***, p<0.001 for indicated virus versus all other conditions at 12 hpi. £, p<0.05 for
947 O/03-K186E-230 versus 230; ££, p<0.01 for O/03-K186E-230 versus O/03-K186E;
948 £££, p<0.0001 for O/03-K186E-230 versus O/03 and mock. +, p<0.05 for indicated
949 virus versus mock-infected samples at 24 hpi. Gamma tubulin was used as a loading
950 control.
951
952 **Figure 6. Comparison of growth kinetics and production of antiviral cytokines**
953 **in equine cells infected with EIV O/03 and NS1 revertant viruses.**
954 E-derm cells were infected (MOI 0.1) for 72h with **(A)** O/03, **(B)** O/03-K186E, **(C)**
955 O/03-230, and **(D)** O/03-K186E-230. Supernatants were collected and cells fixed at
956 indicated times pi. The total amount of antiviral cytokines produced upon infection
957 was measured using an IFN bioassay, and virus growth kinetics was determined by
958 immunofocus assay. Error bars represent SEM. Significance was calculated as
959 indicated in the Method section. ****, p<0.0001 for indicated virus at indicated time
960 post infection versus the three other viruses at the same time post infection.
961
962 **Figure 7. JAK1/2 inhibition restores growth kinetics of NS1 revertant viruses.**
963 E-derm cells were treated for 24h with uIFN **(A)** or Ruxolitinib **(B)** prior to infection
964 (MOI 0.1) with O/03, O/03-K186E, O/03-230, and O/03-K186E-230. Cells and
965 supernatants were collected at different times post-infection. Viral titers were
966 measured as described in Methods. Significance was calculated as indicated in
967 material and method. ***, p<0.001 and *, p<0.05 for O/03 at indicated time post
968 infection versus the three other viruses at the same time post infection.
969

970 **Figure 8. NS1 amino acid 186 and the C-terminal tail affect EIV control of gene**
971 **expression in equine cells.**
972 **(A)** Heat map of differentially expressed genes (DEGs) between O/03 and revertant
973 virus infected-E-derm cells versus mock-infected samples. |Log2FC|>0.58 and
974 p<0.05 are regarded as statistically differentially expressed to mock. The total
975 number of DEGs for each condition is indicated at the bottom of the heat map.
976 DEGs to mock infected conditions were classified in 11 groups: ***1-***DEGs shared
977 between O/03 and mutants; ***2-***DEGs shared between all viruses but O/03-230; ***3-***
978 DEGs shared between all viruses but O/03-K186E; ***4-***DEGs shared between O/03
979 and O/03-K186E; ***5-***DEGs shared between O/03 and O/03-230; ***6-***DEGs shared
980 between O/03 and O/03-K186E-230; ***7-***DEGs shared between all revertants; ***8-***DEGs
981 present in O/03 only; ***9-***DEGs present in O/03-K186E only; ***10-***DEGs present in
982 O/03-230 only; and ***11-***DEGs present in O/03-K186E-230 only. The complete list of
983 genes in presented in Table S1. **(B)** List of DEGs (including some ISGs) present only
984 in revertant virus-infected samples (*group 7*). Values indicate Log2FC.
985 |Log2FC|>0.58 and p<0.05 are regarded as statistically significant.
986
987 **Figure 9. STAT1 nuclear localization in equine cells infected with EIV O/03 and**
988 **NS1 revertant viruses.**
989 E-derm cells were infected (MOI 0.1) with O/03, O/03-K186E, O/03-230, and O/03-
990 K186E-230 viruses or mock infected for 8 or 24h. **(A)** Top: Confocal images showing
991 co-immuno-staining of viral NP and STAT1. The experiment was performed three
992 times independently, and representative pictures of STAT1 and NP are shown. Scale
993 represents 20μm. Bottom: Graphical representation of nuclear localization of STAT1
994 (nSTAT1). nSTAT1 was quantified as described in materials and methods and

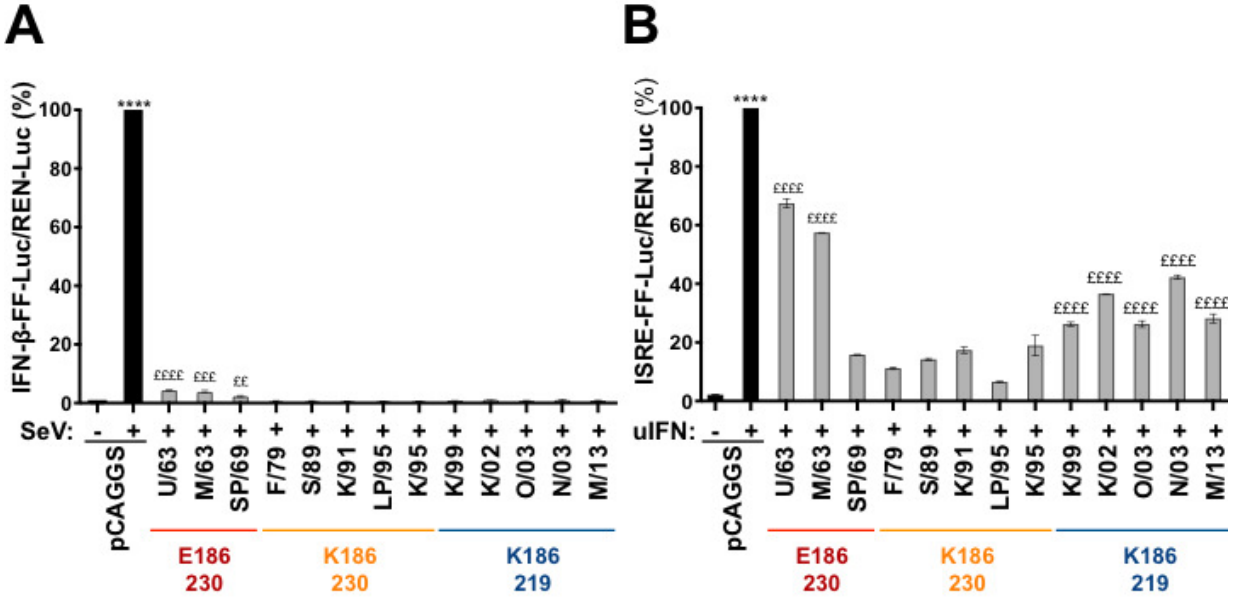
995 expressed as a percentage of total STAT1 (tSTAT1). Statistical significance was
996 calculated as described in materials and methods. **, p<0.01 for O/03-K186E versus
997 other conditions at 8 hpi. No other significant differences were detected. **(B)** Infected
998 cells were lysed and total STAT1 expression was determined by western blot.
999 Gamma tubulin was used as a loading control.
1000
1001 **Figure 10. K186 and C-terminal truncation were fixed at the virus population**
1002 **level during EIV NS1 evolutionary history.**
1003 **(A)** Phylogenetic tree highlighting the fixation of E186K substitution and C-terminal
1004 truncation. Red branches represent NS1 proteins containing E186 and 230 amino
1005 acids (Early NS1), and orange branches represent NS1 proteins containing K186
1006 and 230 amino acids (Intermediate NS1), and blue branches represent NS1 proteins
1007 containing K186 and 219 amino acids (Late NS1).
1008
1009
1010
1011 **TABLES**
1012 **Table 1.** Viruses used in this study.

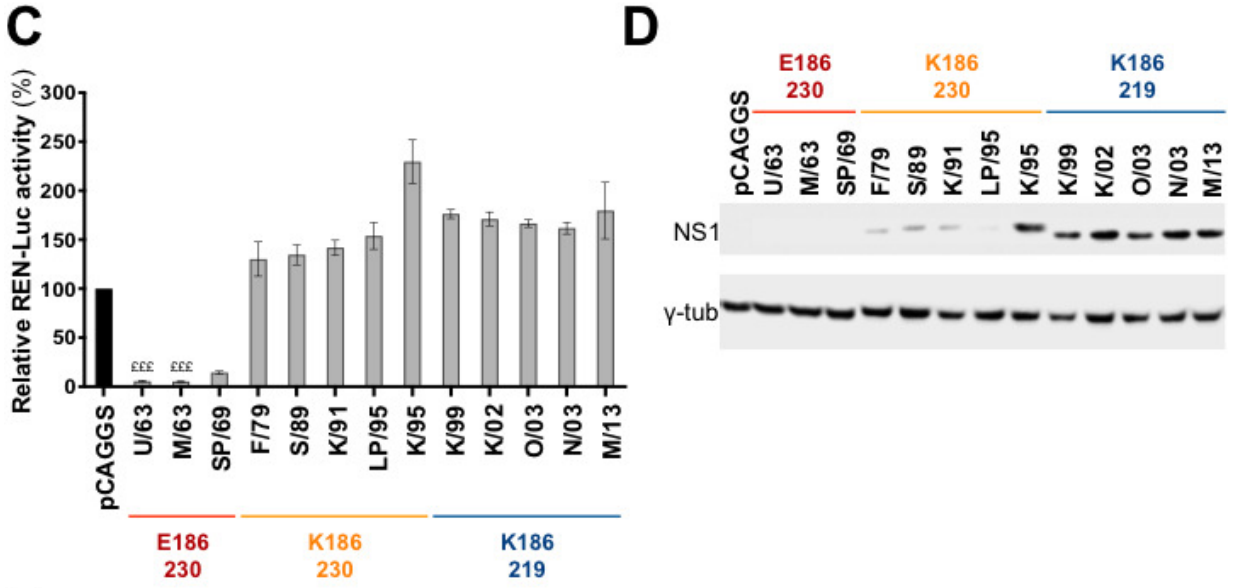
Virus name	Accession number	Abbreviation
A/equine/Uruguay/1/1963	ACD85423	U/63
A/equine/Miami/1/1963	ABY81497	M/63
A/equine/SaoPaulo/1/1969	ACD85390	SP/69
A/equine/Fontainebleau/1/1979	ACD85401	F/79
A/equine/Sussex/1/1989	ACD97430	S/89

A/equine/Kentucky/1/1991	ACA24672	K/91
A/equine/LaPlata/1995	MF182460	LP/95
A/equine/Kentucky/1995	MF182451	K/95
A/equine/Kentucky/1999	MF182443	K/99
A/equine/Kentucky/5/2002	ABA42429	K/02
A/equine/Newmarket/5/2003	ACI48802	N/03
A/equine/Ohio/1/2003	ABA42431	O/03
A/equine/Mongolia/3/2013	MF182459	M/13

1013

Figure 1





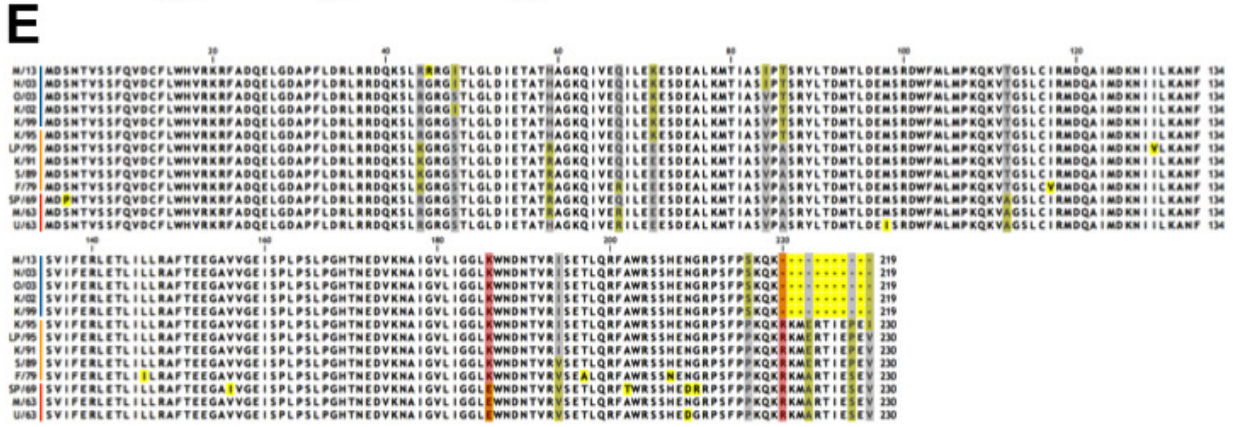
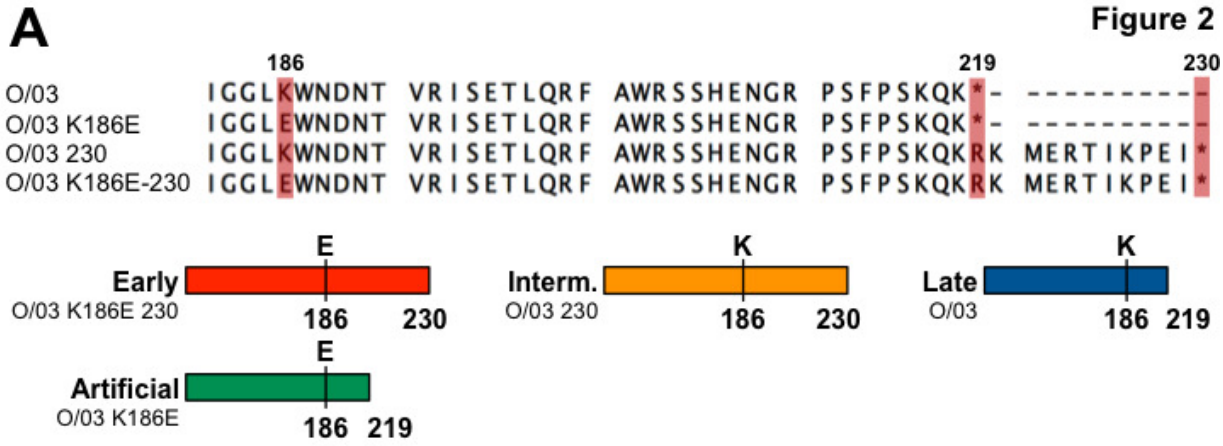
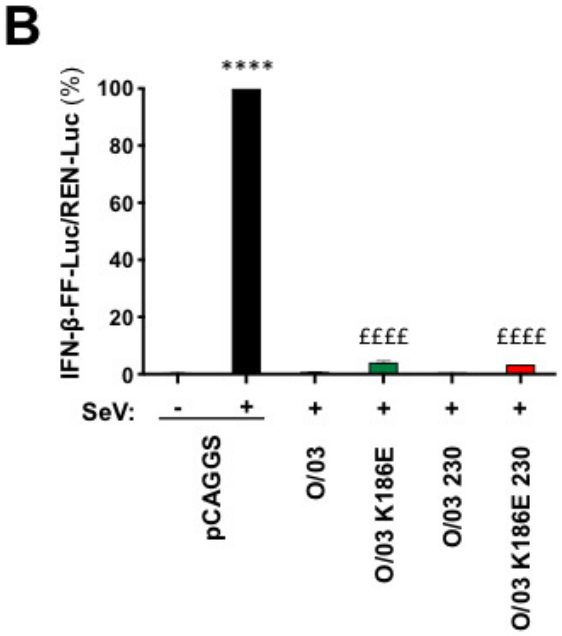
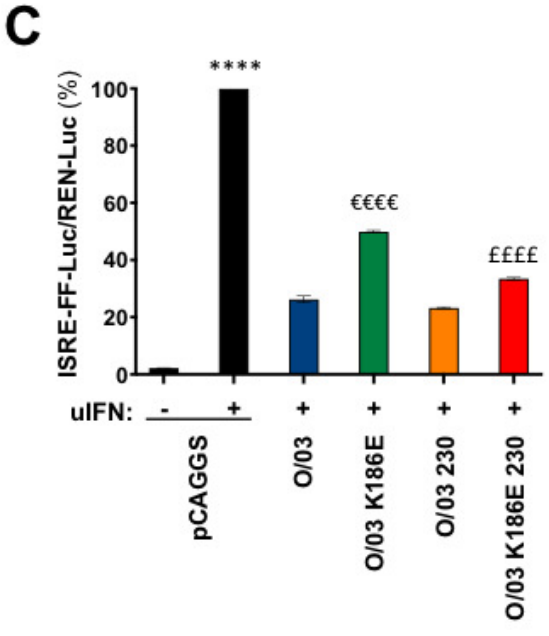
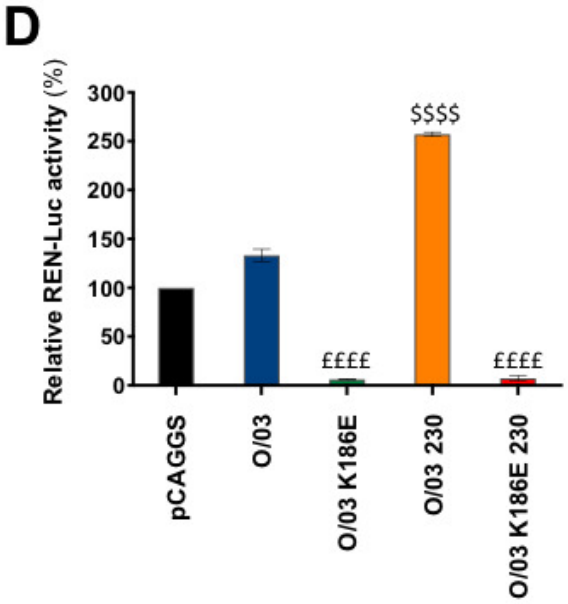


Figure 2









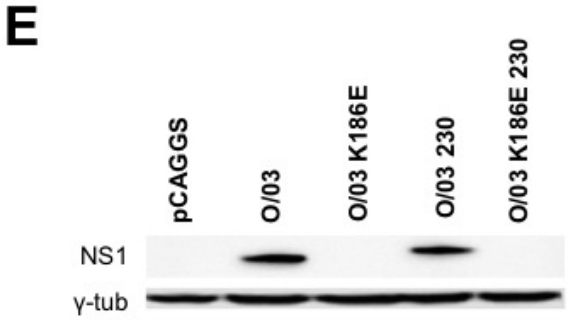


Figure 3

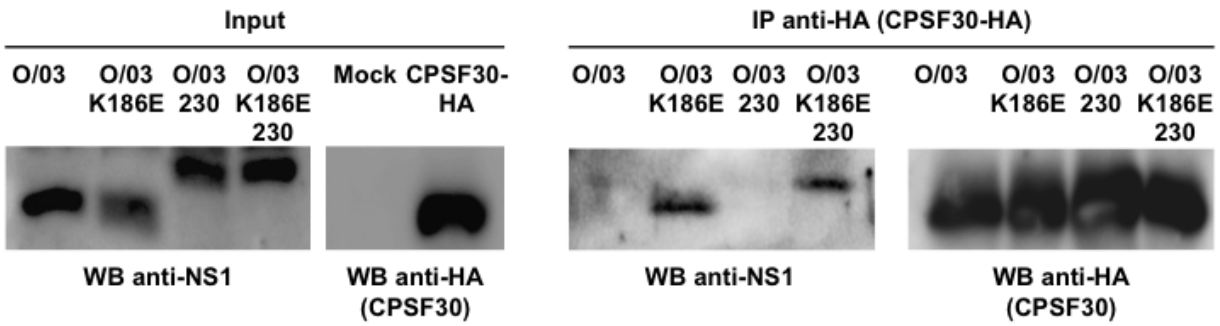


Figure 4

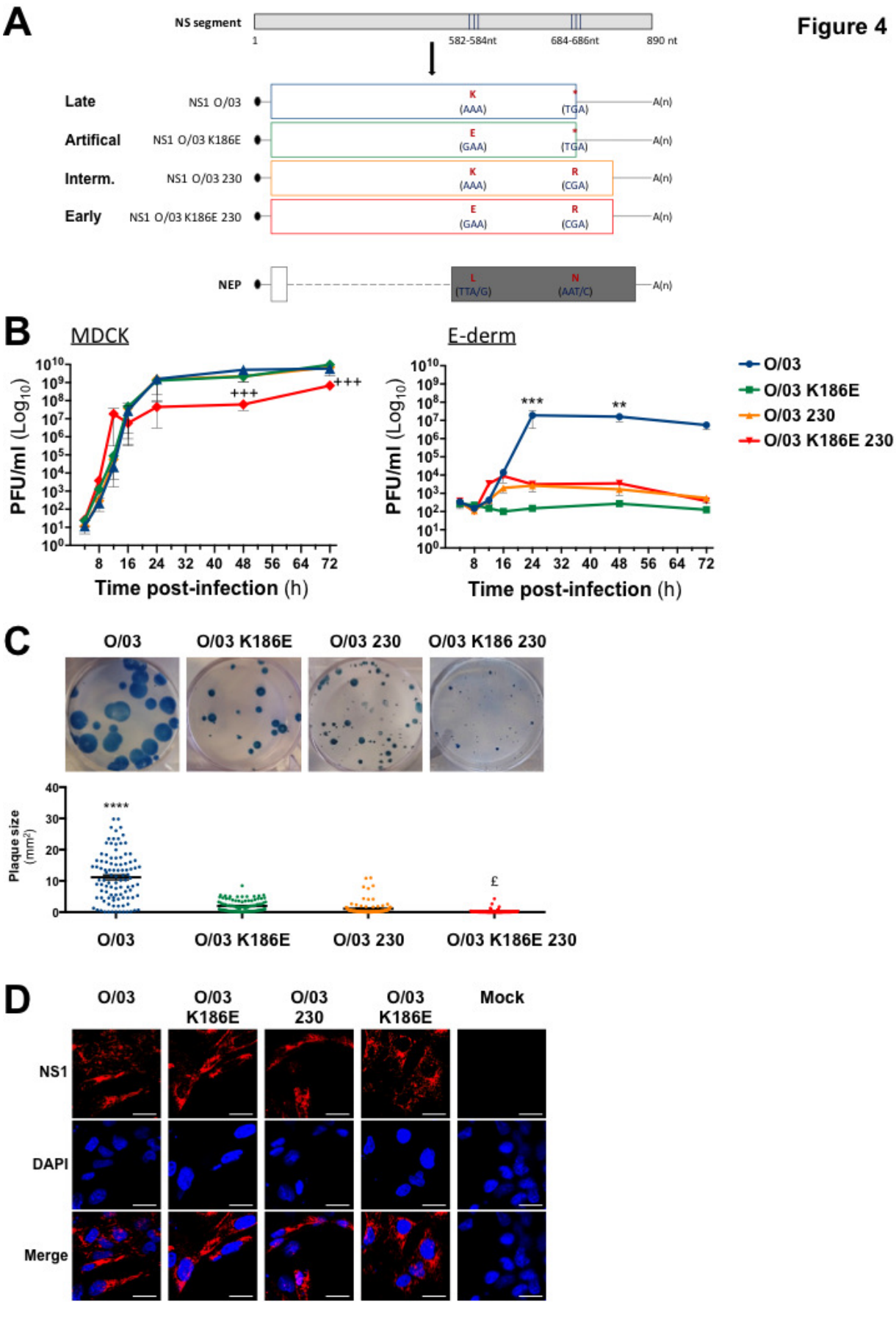


Figure 5

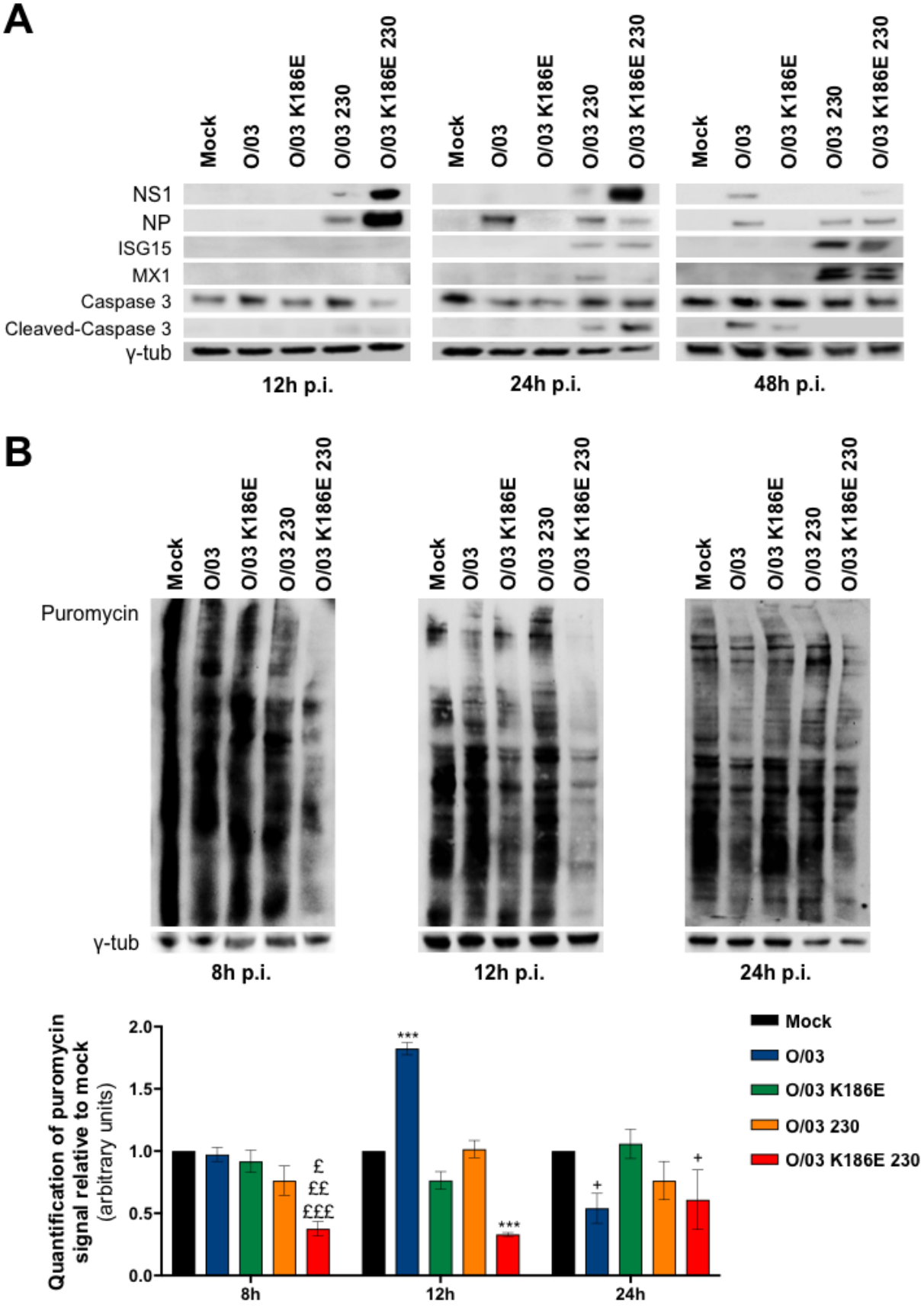


Figure 6

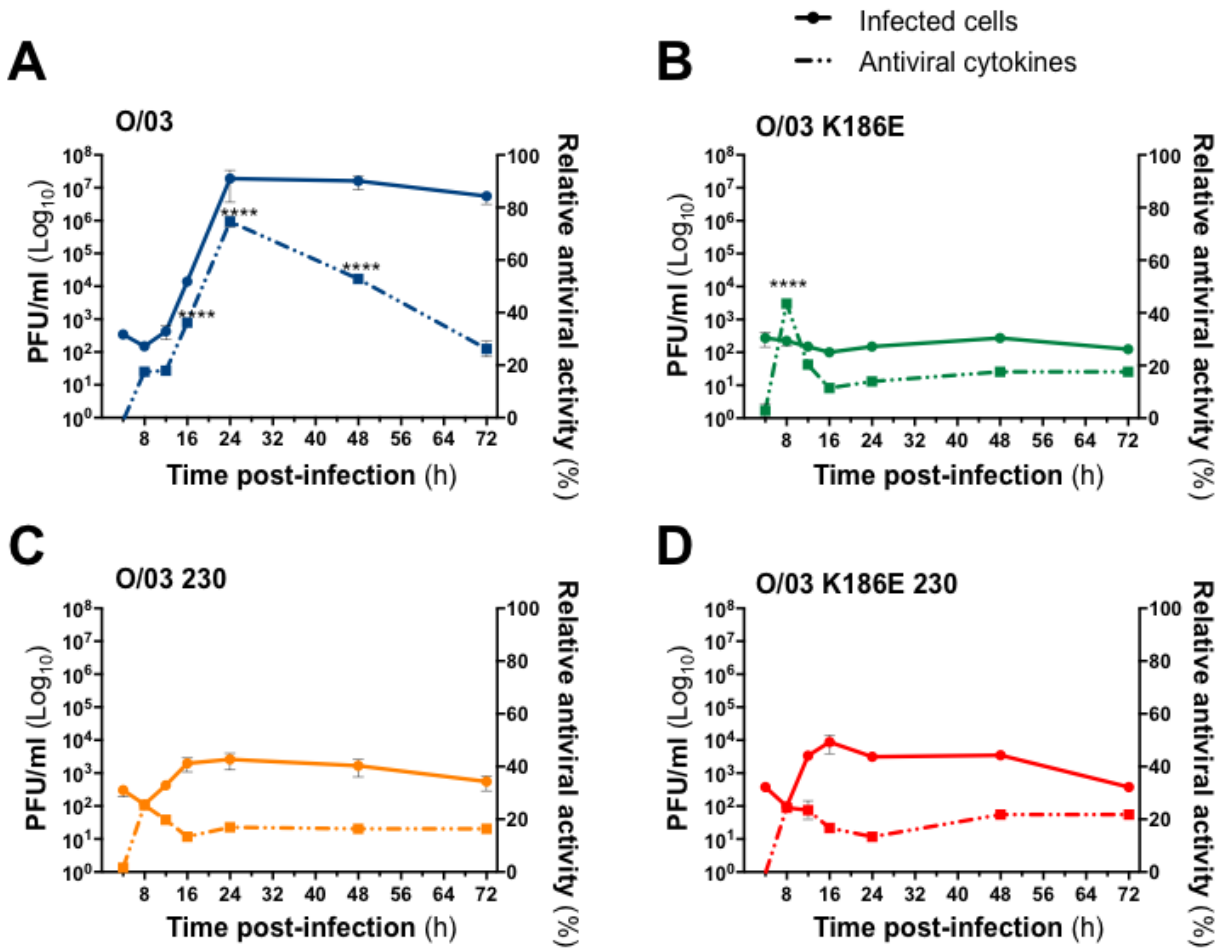


Figure 7

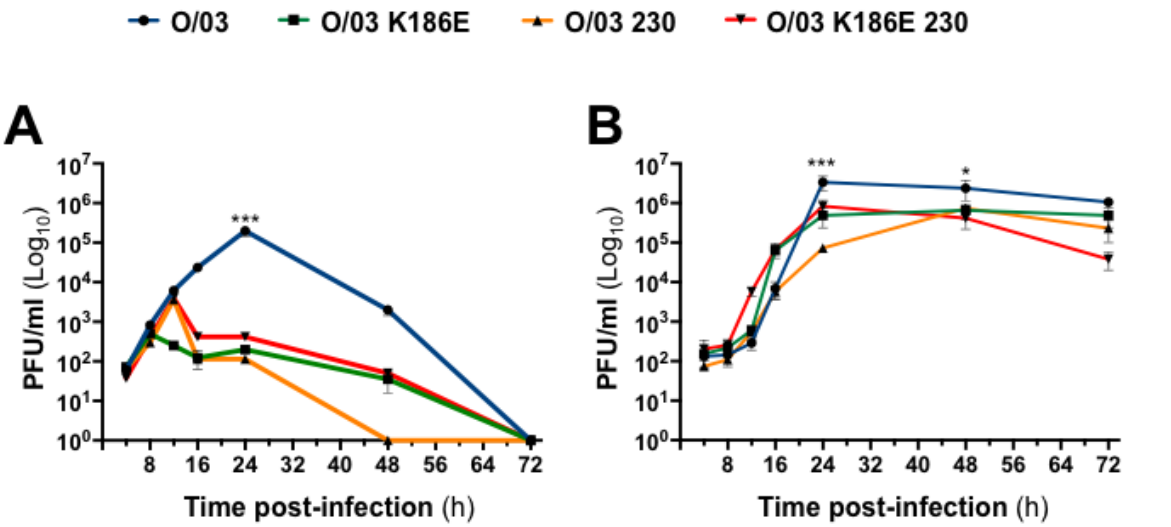


Figure 8

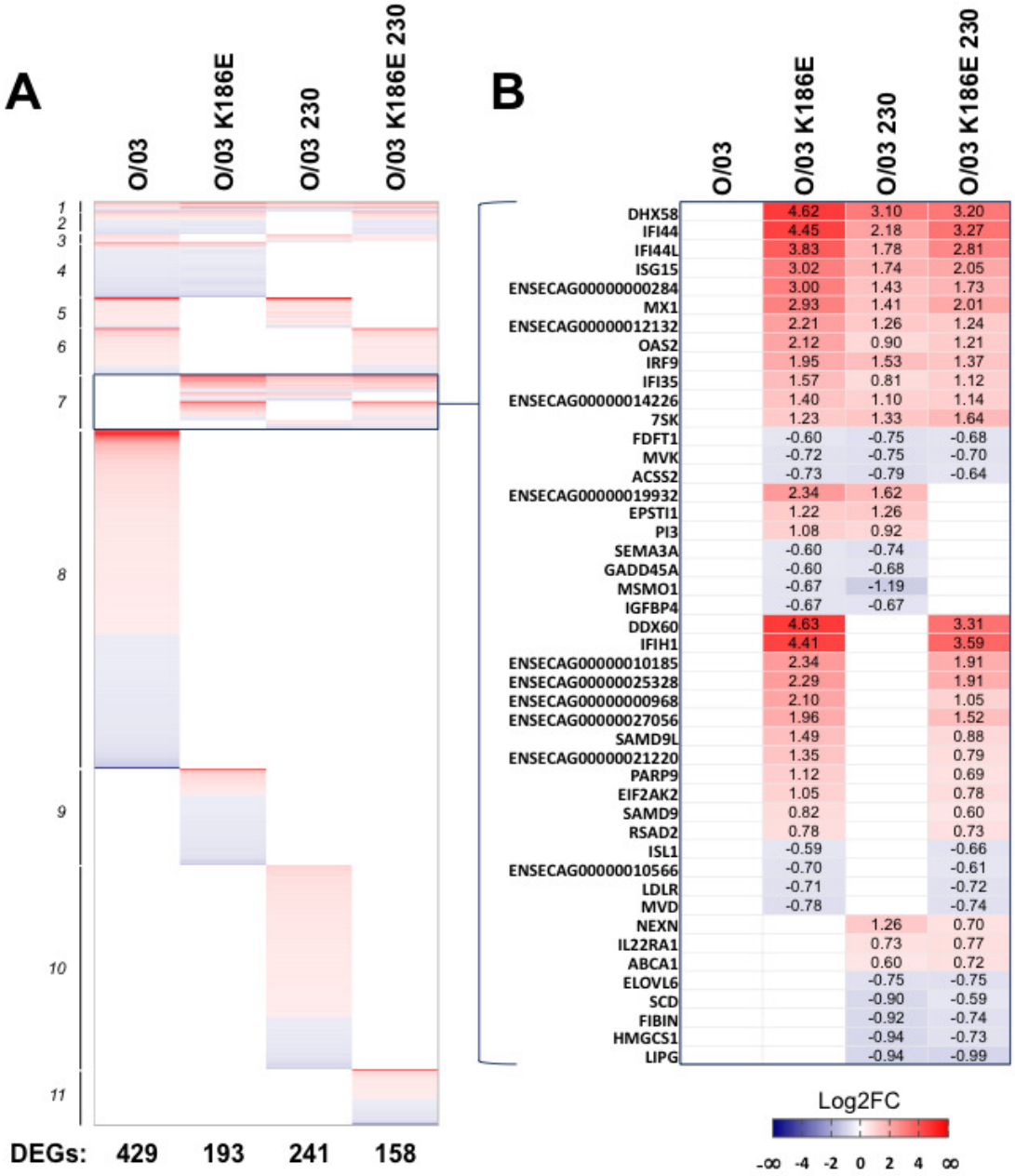


Figure 9

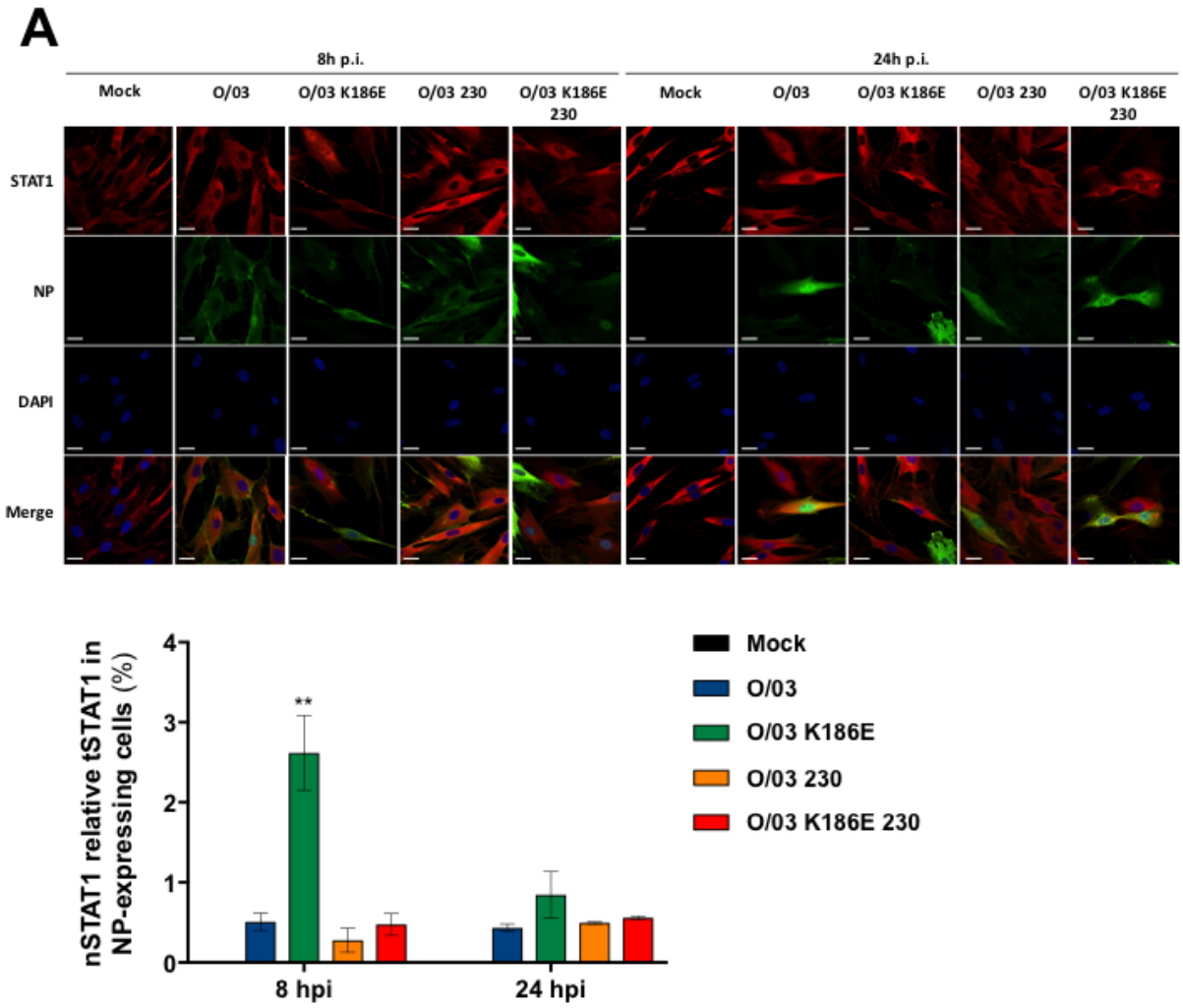


Figure 10